

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Nejc Župec

Klasifikacija Wi-Fi signalov in merjenje obiska po regijah

MAGISTRSKO DELO
ŠTUDIJSKI PROGRAM DRUGE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Mojca Ciglarič

Ljubljana, 2015

Rezultati magistrskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov magistrskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

IZJAVA O AVTORSTVU MAGISTRSKEGA DELA

Spodaj podpisani Nejc Župec sem avtor magistrskega dela z naslovom:

Klasifikacija Wi-Fi signalov in merjenje obiska po regijah

S svojim podpisom zagotavljam, da:

- sem magistrsko delo izdelal samostojno pod mentorstvom doc. dr. Mojce Ciglarič,
- so elektronska oblika magistrskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko magistrskega dela,
- soglašam z javno objavo elektronske oblike magistrskega dela v zbirki "Dela FRI".

V Ljubljani, 24. junija 2015

Podpis avtorja:

Zahvaljujem se mentorici, doc. dr. Mojci Ciglarič, za strokovno pomoč in napotke, ki so mi olajšali pisanje magistrske naloge. Zahvaljujem se staršema in bratu Janu, za podporo in vztrajanje, da čim hitreje zaključim z magistrskim študijem.

Maja, hvala ti za pomoč pri oblikovanju stavkov. Hvaležen sem tudi za spodbudne besede, ki so prišle takrat, ko je bilo to najbolj potrebno.

Jaka, hvala za svetovanje na področju strojnega učenja in prevajanju angleških izrazov.

Hvala tudi ostalim, ki ste mi pomagali in stali ob strani.

Kazalo

1	Uvod	1
1.1	Sestava magistrske naloge	2
2	Wi-Fi	3
2.1	Osnovni pojmi	5
2.2	Prednosti	9
3	Pregled metod za določanje lokacije	11
3.1	Pregled znanstvenih del	12
3.2	Ocena razdalje in kvalitete signala	13
3.2.1	Model pojenjanja signala	16
3.2.2	Kvaliteta signala	19
3.2.3	Primeri izračunov	20
3.3	Problem določanja lokacije	21
3.4	Metoda najmočnejše dostopne točke	22
3.5	Profiliranje signalov Wi-Fi	23
3.6	Trilateracija	25
3.7	Primerjava	27
4	Pregled produktov za določanje lokacije	31
4.1	Cisco MSE	31
4.2	Meraki Presence Analytics	33
4.3	Ruckus SPoT	35
4.4	Primerjava	36

5	Razvoj sistema za zajem signalov Wi-Fi	37
5.1	Arhitektura sistema	37
5.1.1	Dostopne točke	38
5.1.2	Strežnik	41
5.2	Predstavitev podatkov	42
5.2.1	Intervali prisotnosti	42
5.2.2	Vektorski opis signala	43
5.3	Zajemanje signala Wi-Fi z orodjem Wireshark	44
5.4	Problem zasebnosti	47
6	Klasifikacija signalov Wi-Fi v regije	51
6.1	Definicija problema	52
6.2	Klasifikacija z metodo najmočnejše dostopne točke	54
6.3	Klasifikacija s trilateracijo	55
6.4	Klasifikacija s strojnim učenjem	57
6.5	Od klasifikacije do obiska po regijah	59
7	Postavitev sistema in statistika zajetih podatkov	61
7.1	Shema postavitve	62
7.2	Statistika zajetih podatkov	64
7.2.1	Statistika intervalov prisotnosti	66
7.2.2	Statistika vektorjev	67
8	Analiza rezultatov	71
8.1	Posamezna dostopna točka	71
8.1.1	Obisk skozi čas	72
8.1.2	Obiskanost po dnevih v tednu	76
8.1.3	Časi zadrževanja	77
8.1.4	Korelacija obiska in vremena	79
8.2	Več dostopnih točk	80
8.2.1	Zaznavanje predavanj in odmorov	81

KAZALO

9	Evalvacija klasifikacije signalov Wi-Fi	85
9.1	Naključno razporejanje	85
9.2	Metoda najmočnejše dostopne točke	86
9.3	Trilateracija	87
9.4	Klasifikacija s strojnim učenjem	88
10	Sklepne ugotovitve	93
10.1	Nadaljnje delo	95
A	Spletni portal <i>Wi-Fi Analytics</i>	103
B	Vreme in obisk po dnevih	111
C	Meritve signalov po regijah	115

Seznam uporabljenih kratic

kratica	angleško	slovensko
AP	access point	dostopna točka
BLE	Bluetooth low energy	Bluetooth z nizko porabo energije
BSS	basic service set	osnovna enota brezžičnega omrežja
GPS	global positioning system	globalni sistem pozicioniranja
IPS	indoor positioning system	sistem pozicioniranja v notranjih prostorih
PCA	principal component analysis	metoda glavnih komponent
RSSI	received signal strength indication	indikator moči sprejetega signala
RTLS	real-time locating system	lokacijski sistem v realnem času
SSID	service set identifier	ime brezžičnega omrežja
SVM	support vector machine	metoda podpornih vektorjev
WEP	wire equivalent privacy	zasebnost enakovredna žičnemu prenosu
WLAN	wireless local area network	brezžično lokalno omrežje
WPA	Wi-Fi protected access	Wi-Fi zaščiten dostop
SNR	signal-to-noise ratio	razmerje med signalom in šumom

Povzetek

Vse več je mobilnih naprav, ki oddajajo signale Wi-Fi. Te lahko zaznamo z dostopnimi točkami, posledično pa lahko merimo število obiskovalcev v bližini dostopnih točk. Pogosto nas zanima število obiskovalcev samo za določen prostor oz. regijo. Zato smo v okviru magistrske naloge raziskali metode in produkte, s katerimi je možno določiti lokacijo izvora signala. Ker nas zanima le regija, smo problem določanja lokacije transformirali v problem klasifikacije signala Wi-Fi v regije. Po zgledu obstoječih metod za določanje lokacije (metoda najmočnejša dostopne točke, trilateracija in sistem RADAR), smo razvili tri nove metode. Da bi lahko te metode ovrednotili, smo na Fakulteti za računalništvo in informatiko postavili sistem za zajem signalov Wi-Fi in z njim dva meseca zajemali podatke. S pomočjo strojnega učenja smo razvili metodo, ki v 85,1 % pravilno napove regijo izvora signala Wi-Fi. V primeru, da so regije med seboj ločene s stenami, pa klasifikacijska natančnost znaša več kot 93 %.

Ključne besede: Wi-Fi, klasifikacija, SVM, strojno učenje, RADAR, trilateracija, določanje lokacije

Abstract

More and more mobile devices transmit Wi-Fi signals which can be detected by access points. Consequently the number of visitors near access points can be measured. Often we are interested in the number of visitors only for a certain region. Therefore, in the context of master's thesis, methods and products for indoor localization were studied. We transformed the localization problem into classification of Wi-Fi signals problem. Based on existing methods (nearest base station, trilateration and RADAR) we developed three new methods. In order to evaluate these methods the system for capturing Wi-Fi signals was set up at the Faculty of Computer and Information Science in Ljubljana. The data was being collected for two months. Based on machine learning algorithms we developed a new method, which correctly predicts the region in 85,1 % cases. If regions are separated by walls, the classification accuracy is more than 93 %.

Keywords: Wi-Fi, classification, SVM, machine learning, RADAR, trilateration, localization

Poglavje 1

Uvod

Vse več je pametnih mobilnih naprav, ki ves čas oddajajo signale Wi-Fi. Ker lahko te signale zaznamo z dostopnimi točkami, lahko ugotovimo, koliko naprav se nahaja v bližini dostopnih točk. Na ta način lahko merimo obiskanost posameznih področij v notranjih prostorih. Na trgu obstajajo številni produkti, ki omogočajo določiti lokacijo mobilnih naprav, vendar pa so ti praviloma zelo dragi, hkrati pa zahtevajo veliko časa za vzpostavitev sistema, predvsem na račun kompleksne kalibracije.

V okviru magistrske naloge bomo raziskali najboljše metode, za določanje lokacije v notranjih prostorih. Na podlagi tega znanja, bomo poizkušali poiskati kompromis med zahtevnostjo vzpostavitve sistema in natančnostjo. Zgradili bomo nov sistem, ki bo omogočal zajem signalov Wi-Fi, hkrati pa bo enostaven in ne bo zahteval veliko časa za kalibracijo. Pri tem bomo uporabili poceni računalnike Raspberry Pi. Sistem bomo nato postavili na realno lokacijo, kjer bomo nekaj časa zbirali podatke, nato pa jih bomo natančno analizirali in preverili, ali se da s preprostim zajemom signala Wi-Fi določiti dejanski obisk. V nadaljevanju bomo šli še en korak dlje in bomo poskušali ugotoviti regijo izvora signala. Zato bomo razvili metode, ki omogočajo klasifikacijo signalov po regijah. Ob koncu bomo te metode še evalvirali in preverili, kako se obnesejo v praksi.

Naše metode bodo temeljile na zajemanju prometa Wi-Fi. Pri tem bomo

zajemali tudi naslove MAC, ki veljajo za občutljive podatke. Zato bomo poiskali način, kako zagotoviti čim višjo stopnjo zasebnosti.

Poleg signalov Wi-Fi, bomo zajemali tudi vremenske podatke, na podlagi katerih bomo poskušali najti korelacijo med obiskom in vremenom.

1.1 Sestava magistrske naloge

V drugem poglavju podrobneje predstavimo tehnologijo Wi-Fi. Opišemo protokol, ki omogoča povezovanje mobilnih naprav z brezžičnimi omrežji in razložimo, zakaj je možno pasivno spremljanje mobilnih naprav. Ob koncu poglavja predstavimo glavne prednosti tehnologije Wi-Fi. Tretje poglavje je namenjeno teoretični predstavitvi metod za določanje lokacije v notranjih prostorih. Razložimo tudi, kako je možno na podlagi moči signala oceniti razdaljo. V četrtem poglavju pregledamo tri komercialne produkte, ki ponujajo rešitev za določanje lokacije v notranjih prostorih. V petem poglavju predstavimo razvoj novega sistema za zajem signalov Wi-Fi, ki bazira na poceni računalnikih Raspberry Pi. Opišemo glavne pasti, na katere smo naleteli in kako smo jih rešili. V šestem poglavju uvedemo problem klasificiranja signalov med različnimi regijami. Predstavimo tudi tri metode, s katerimi lahko rešimo ta problem. V sedmem poglavju opišemo praktično testiranje sistema za zajem podatkov Wi-Fi in statistično analiziramo zajete podatke. Osmo poglavje je namenjeno predstavitvi konkretnih rezultatov. Podrobneje predstavimo zaznan obisk na Fakulteti za računalništvo in informatiko. V devetem poglavju ocenimo natančnost metod za klasifikacijo signalov Wi-Fi, v desetem pa predstavimo sklepne ugotovitve. V dodatku A na kratko predstavimo spletni portal *Wi-Fi Analytics*. V dodatku B so podrobneje predstavljene meritve obiska po dnevih. V dodatku C se nahajajo meritve signalov po regijah.

Poglavje 2

Wi-Fi

Wi-Fi je tehnologija, ki omogoča brezžično povezljivost med elektronskimi napravami in lokalnimi omrežji LAN (local area networks). Za njo stoji organizacija Wi-Fi Alliance, ki skrbi za standarde. Čeprav je imel standard sprva nekaj težav pri zagotavljanju varne komunikacije¹, lahko danes govorimo o zreli in varni tehnologiji, ki se je zaradi svoje svoje praktičnosti v zadnjem desetletju močno razširila. Danes se čipi Wi-Fi vgrajujejo v prenosne računalnike, tablice, mobilne telefone, igralne konzole, kamere, fotoaparate, itd. Te naprave so postale zelo praktične, saj ni več potrebno iskati žične povezave. V prihodnosti lahko tako pričakujemo še več naprav, ki bodo uporabljale brezžično povezljivost, posledično pa bo na voljo tudi več brezžičnih omrežij. Nekaj dejstev, ki govorijo temu v prid:

Razbremenitev mobilnih omrežij (angl. mobile data offloading) - Mobilni operaterji se soočajo z naraščanjem mobilnega podatkovnega prenosa in iščejo načine, kako zagotoviti čim boljše uporabniško izkušnjo in cenovno ugodno, razširljivo ter dobičkonosno rešitev. Statistike kažejo,

¹Wi-Fi paketi se prenašajo po zraku, ki jih lahko prebere vsak, ki je dovolj blizu. Zato je potrebno te pakete zašifrirati. Prva splošna enkripcija WEP (angl. wired equivalent privacy) se je kmalu izkazala za ranljivo in danes ne zagotavlja anonimne komunikacije [36]. Nova načina enkripcije (WPA oz. WPA2) sta problem slabe enkripcije rešila in tako lahko govorimo o varni brezžični povezljivosti.

da se mobilni uporabniki večino časa (več kot 80 % [23, 20]) zadržujejo v notranjih prostorih, torej v pisarnah, domovih, nakupovalnih središčih, postajah javnega prometa, itn. Uporabniki so se v teh prostorih navadili uporabljati Wi-Fi, saj je cenejši od mobilnih podatkov (LTE, 3G, Edge). Na ceno vplivajo tudi koncesije za frekvenčni spekter, ki jih morajo operaterji plačevati državam. Po drugi strani pa Wi-Fi dela na frekvenčnem spektru, ki je brezplačen. Operaterji zato vidijo Wi-Fi kot poceni razširitev njihovega omrežja, s katerim bi uporabnikom omogočili cenejši dostop do interneta, sami pa bi s tem preprečili izpad zaslužka. Na podlagi navedenih dejstev lahko sklepamo, da se bo postavilo še več omrežij Wi-Fi, uporabniki mobilnih naprav pa bodo imeli še pogosteje vključeno Wi-Fi povezljivost.

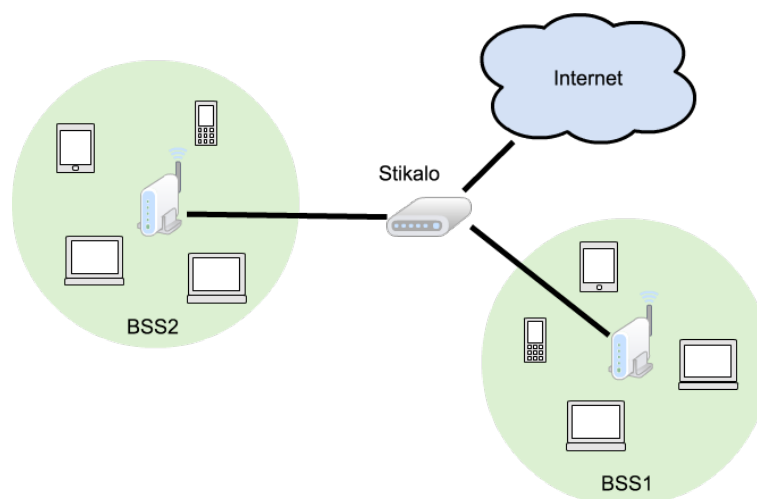
Internet stvari (angl. Internet of Things) - Vse več različnih naprav se vključuje na internet, med drugim tudi hladilnike, televizije, pečice, sobne termostate, pametne ure, pametne žarnice, itd. Govorimo o t.i. internetu stvari, ki med seboj povezuje vse več elektronskih naprav in omogoča nove storitve, ki nudijo udobnejše življenje. Te naprave za povezovanje uporabljajo med drugim tudi Wi-Fi, vse bolj popularna pa postajata tudi Bluetooth Low Energy (BLE) in ZigBee, ki omogočata brezžično povezljivost napravam z majhnimi baterijami.

Gosto naseljena območja - V gosto naseljenih predelih se pogosto pojavi problem zasičenosti frekvenc. Podobno se dogaja na množičnih prireditvah, npr. na koncertih oz. športnih prireditvah. Ljudje danes vse pogosteje uporabljajo mobilne naprave, s katerimi slikajo/snemajo dogajanje, hkrati pa želijo te vsebine čim hitreje prenesti na socialna omrežja, kjer svoja doživljanja delijo s svojimi prijatelji. Mobilna omrežja niso grajena za tako veliko množico mobilnih naprav in razbremenitev mobilnih frekvenc je nujno potrebno. Nekatera mesta postavljajo omrežja Wi-Fi, ki predvsem turistom omogočajo poceni dostop do spletnih storitev, npr. Ljubljana je postavila omrežje WiFree [34].

Z razvojem pametnih mest (angl. smart cities) bo takih omrežij v prihodnosti vse več. Med drugim tudi Evropska unija ponuja finančna sredstva [5], s katerimi želi vzpodbuditi razvoj infrastrukture, da bo lahko kos vse večji potrebi po širokopasovnem dostopu mobilnih naprav.

Vsi ti našteti vzroki so privedli do tega, da je Wi-Fi infrastruktura že danes močno razširjena. V bodoče pa lahko pričakujemo še več omrežij Wi-Fi, predvsem na račun hitro-razvijajočih se držav.

2.1 Osnovni pojmi

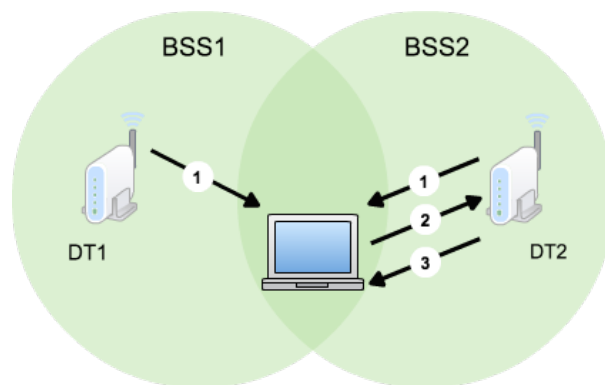


Slika 2.1: Primer dveh brezžičnih omrežij, ki sta preko stikala/usmerjevalnika povezani v internet.

Osnovna enota (angl. basic station set, BSS) vsakega brezžičnega omrežja je sestavljena iz dostopne točke (angl. access point) in ene ali več brezžičnih postaj (angl. wireless stations). Vsako brezžično omrežje se predstavlja s t.i. SSID-jem (angl. Service Set Identifier). Na sliki 2.1 lahko vidimo dve osnovni enoti, ki sta preko stikala/usmerjevalnika povezani v internet [25].

Vsaka dostopna točka/brezžična postaja ima v strojni programske opremi (angl. firmware) zapisan naslov MAC. Sestavljen je iz šestih bajtov oz. 48

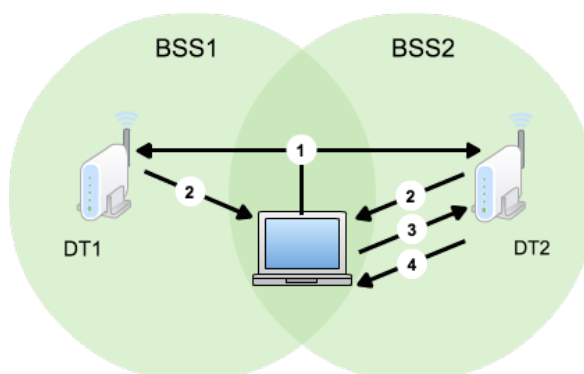
bitov. Prvih 24 bitov označuje proizvajalca čipa, naslednjih 24 bitov pa mora proizvajalec porazdeliti tako, da vsaka naprava prejme unikatni naslov. Posledično se iz naslovov MAC lahko razbere proizvajalca določenega čipa. Organizacija IEEE ves čas posodablja bazo naslovov MAC, ki jih dodelijo proizvajalcem mrežnih čipov [7]. Na ta način lahko zelo enostavno ugotovimo, ali se v omrežje povezuje telefon znamke Apple ali pa gre za telefon, ki je bil narejen pri Samsungu. V teoriji naj bi bili vsi naslovi MAC enolični, a v praksi ni vedno tako, saj je naslov MAC možno spremeniti tudi s programsko opremo.



Slika 2.2: Pasivno skeniranje: (1) dostopni točki pošljeta *beacon* okvir, (2) računalnik pošlje *association request* okvir do dostopne točke, (3) dostopna točka pošlje *association response* okvir nazaj do računalnika.

Ko želi brezžična postaja vzpostaviti povezavo z internetom, mora najprej poiskati vsa brezžična omrežja, ki so na voljo v okolici in za katere pozna ključne. Te išče s pasivnim oz. aktivnim skeniranjem:

Pasivno skeniranje - Standard 802.11 zahteva [8], da dostopna točka periodično pošilja okvirje (angl. frames), ki vsebujejo SSID in naslov MAC dostopne točke. Brezžične postaje morajo ves čas poslušati na vseh enajstih kanalih. Po določenem času ima naprava na voljo seznam vseh razpoložljivih omrežij v okolici. Glavna prednost pasivnega skeniranja je manjša poraba energije, saj napravi ni potrebno pošiljati okvirjev, temveč samo poslušati, kar zahteva manj energije.



Slika 2.3: Aktivno skeniranje: (1) računalnik pošlje broadcast *probe request*, (2) dostopni točki odgovorita s *probe response*, (3) računalnik pošlje *association request* okvir do dostopne točke, (4) dostopna točka pošlje *association response* okvir nazaj do računalnika.

Aktivno skeniranje - Brezžična postaja pošlje *probe request* okvir (v broadcast načinu), ki ga vidijo vse bližnje dostopne točke. Te odgovorijo s *probe response* okvirjem, v katerih se nahaja naslov MAC dostopne točke in SSID. Tako naprava zelo hitro pride do seznama vseh brezžičnih omrežij v okolici. Alternativno lahko postaja pošlje *probe request* (unicast način) s točno določenim SSID-jem in tako preveri, če je specifično brezžično omrežje na voljo. Na ta način lahko naprava odkrije tudi brezžična omrežja, ki ne oglašujejo SSID-ja (tista, ki uporabljajo skrit SSID). Aktivno skeniranje je hitrejše od pasivnega, saj naprava v krajšem času prejme vse okvirje.

Ko naprava pošlje *probe request* na enem kanalu, praviloma počaka 20-40 milisekund, da prispejo vsi odgovori dostopnih točk, nato pa se premakne na naslednji kanal in pošlje naslednji *probe request*. Ker je kanalov enajst, naprava potrebuje manj kot pol sekunde², da pridobi seznam vseh omrežij. Pri pasivnem skeniranju naprava čaka 100 milisekund in več. Aktivno skeniranje je tako lahko tudi do 2,5 krat hitrejše od pasivnega [1].

Mobilne naprave večinoma uporabljajo aktivno skeniranje. Danes je na-

² $11 * 40 \text{ milisekund} = 440 \text{ milisekund}$

mreč zelo pomembna uporabniška izkušnja in ko uporabnik zahteva seznam brezžičnih omrežij, želi, da se le-ta čim hitreje prikaže. Ko ima naprava prižgan Wi-Fi, ves čas sprašuje, ali se lahko priključi na katero izmed omrežij prisotnih v okolici. Vsaka naprava Wi-Fi hrani seznam omrežij, na katera se je v preteklosti že priključila in ima ključe za prijavo. Zato ves čas sprašuje, če je morda kakšno tako omrežje na voljo in če je, vzpostavi povezavo. Ko se vzpostavi povezava, se začnejo prenašati podatkovni paketi.

Nekateri novejši modeli telefonov upoštevajo tudi premike telefona. Te zaznajo s senzorji; lokalne premike zaznajo žiroskopi, večje premike pa modul GPS. Ko se telefon premika, še intenzivneje išče nova omrežja, medtem pa oddaja pakete *probe request*. Meritve kažejo, da naprave oddajajo pakete *probe request* vsakih 15-60 sekund [11], odvisno od proizvajalca naprave.

Vidimo lahko, da naprave po zraku pošiljajo različne pakete. Do zdaj smo spoznali podatkovne pakete in pakete *probe request*. Slednji so zanimivi predvsem zato, ker se oddajajo tudi takrat, ko naprava ni povezana v omrežje. To pomeni, da dostopne točke vidijo vse naprave, ki imajo zgolj aktivirano Wi-Fi povezljivost! V nadaljevanju se bomo zato osredotočili predvsem na pakete **probe request**. Teh paketov je sicer manj, a dovolj za pridobitev vzorca, ki omogoča nadaljnjo analizo. S tem zmanjšamo kompleksnost procesiranja, saj bi obdelovanje vseh paketov, ki se pošiljajo po zraku, zahtevalo ogromno procesorske moči.

Glede na to, da se ti paketi prenašajo po zraku, do katerega ima dostop vsaka naprava, se poraja vprašanje, kako bi lahko analizirali te pakete. Izkaže se, da je dovolj le navadna brezžična kartica, ki se danes nahaja v vsakem prenosnem računalniku, ter ustrezna programska oprema, npr. odprtokodno orodje Wireshark, katero omogoča zajemanje paketov ter nadaljnjo analizo (več o tem v razdelku 5.3).

2.2 Prednosti

Wi-Fi danes predstavlja standarden način brezžičnega povezovanja mobilnih naprav z brezžičnimi omrežji. Na vidiku so nove generacije, ki bodo prinesle še višje hitrosti prenosa podatkov in daljši doseg. Število omrežij hitro raste in prav nič ne kaže, da bi se ta trend obrnil.

V zadnjih nekaj letih se hitro razvijajo nove tehnične rešitve, ki omogočajo določanje lokacije mobilnih naprav na podlagi oddanih signalov Wi-Fi. Vsi večji proizvajalci brezžične opreme (Cisco, Ruckus Wireless, Meraki, Aruba Networks, Ubiquiti, AeroScout) razvijajo produkte in servise, ki omogočajo določitev lokacije mobilne naprave do nekaj metrov natančno. Te rešitve so prvenstveno namenjene navigiranju v notranjih prostorih. Na trg pa prihajajo tudi storitve, ki omogočajo beleženje obiska in zaznavanje gneče.

Lokacijo mobilnih naprav v notranjih prostorih je mogoče določiti tudi z drugimi tehnologijami, npr. Bluetooth ali RFID. V primeru uporabe teh dveh tehnologij je potrebno namestiti ustrezno strojno opremo, ki zaznava signale. Pri Wi-Fi pa so dostopne točke pogosto že nameščene, kar zmanjša stroške vzpostavitve sistema za določanje lokacije.

V tem poglavju smo spoznali, da mobilne naprave oddajajo signale Wi-Fi tudi če niso povezane v brezžično omrežje. To pomeni, da lahko z ustrezno opremo zaznavamo signale vseh mobilnih naprav, ki imajo prižgan vmesnik Wi-Fi. Teh pa je iz dneva v dan več, saj potreba po podatkovnem prenosu raste, Wi-Fi pa ponuja cenovno dostopno rešitev in zaradi česar vse več uporabnikov hodi naokoli z mobilnim telefonom in prižganim vmesnikom Wi-Fi. Ker pa mobilne naprave ves čas iščejo omrežja, pogosto oddajajo signale. To pa nam omogoča, da lahko s pomočjo sistemov za določanje lokacije določimo število obiskovalcev v bližini dostopnih točk in relativni trend obiska skozi čas.

Poglavje 3

Pregled metod za določanje lokacije

Napredek digitalne tehnologije v zadnjih desetletjih je tako rekoč rešil problem določanja lokacije na prostem. V veliki meri je za to zaslužna ameriška vojska, ki je leta 1973 razvila satelitski sistem GPS (angl. Global Positioning System) [28]. Da lahko sprejemnik GPS določi lokacijo, mora sprejemati signale z vsaj štirimi sateliti¹. Ko sprejemnik izmeri razdalje do vidnih satelitov, s pomočjo trilateracije² izračuna lokacijo na Zemlji. Težava pa se pojavi v zaprtih prostorih, saj so iz materialov, ki otežujejo prehod satelitskim signalom [16]. Za določitev 3D lokacije mora sprejemnik v vsakem trenutku sprejemati signale iz vsaj štirih satelitov, a je ta pogoj redko izpolnjen in posledično lociranje v notranjih prostorih z GPS ni mogoče.

V zaprtih prostorih je določanje lokacije precej težje kot na prostem, saj so prisotne številne ovire, ki motijo signal: stene, strehe, okna, izolacija, električne napeljave in ostale ovire. Kompleksnost dvigujejo tudi različne vrste materialov, ki različno prepuščajo signale in še dodatno otežujejo določanje lokacije [40]. Zaradi hitrega razmaha pametnih mobilnih naprav, potreba po

¹Že trije sateliti zadostujejo za izračun 2D lokacije. Za 3D lokacijo pa so potrebni vsaj štirje sateliti. Vsak dodaten satelit izboljša natančnost lokacije.

²Trilateracija je matematični postopek, ki omogoča izračun lokacije na podlagi razdalj do ostalih točk, za katere je lokacija v naprej poznana.

določanju lokacije v notranjih prostorih vse hitreje narašča. V zadnjih letih so se razširile različne metode, ki temeljijo na brezžičnih tehnologijah (Wi-Fi, Bluetooth, RFID, ZigBee, IR), čeprav te prvenstveno niso namenjene določanju lokacije. Na podlagi brezžičnih signalov se določi razdalja med izvorom signala in dostopnimi točkami, za katere poznamo lokacijo. Nato se uporabijo različni algoritmi, s katerimi določimo lokacijo izvora signala.

Problem določanja lokacije v notranjih prostorih je še danes v izziv številnim znanstvenikom po vsem svetu. V naslednjem razdelku bomo zato pregledali glavna dela s tega področja, pri tem pa bomo poudarili raziskave s področja lociranja Wi-Fi. Nato bomo natančneje pregledali, kako dobro je možno oceniti razdaljo med izvorom signala in dostopno točko, na podlagi moči signala. Če bi lahko to razdaljo natančno izmerili, potem bi problem določanja lokacije rešili na enak način, kot pri GPS, torej s trilateracijo, ki jo bomo podrobneje predstavili ob koncu poglavja. Raziskali bomo še trenutno najpogostejši način za določanje lokacije, to je profiliranje signala Wi-Fi (angl. finger-printing). Na koncu bomo metode še primerjali med seboj, predvsem s stališča zahtevnosti postavitve.

3.1 Pregled znanstvenih del

V zadnjih dveh desetletjih so številni znanstveniki po vsem svetu veliko truda namenili iskanju najboljših metod za določanje lokacije v notranjih prostorih. Že v začetku devetdesetih so dr. Want in kolegi razvili *The Active Badge Location System* [37], ki temelji na infrardeči (IR) svetlobi. Glavni problem sistema je zelo kratek doseg infrardeče svetlobe, posledično pa tudi cena, saj sistem potrebuje veliko sprejemnikov IR. A v tistem času primernejše tehnologije od IR še niso poznali.

Proti koncu tisočletja so se zelo hitro širila omrežja Wi-Fi. Dr. Bahl in dr. Padmanabhan sta leta 2000 v raziskovalnem oddelku korporacije Microsoft, razvila RADAR [9], ki omogoča določanje lokacije na podlagi signalov Wi-Fi in obljublja povprečno natančnost do treh metrov, kar še danes velja

za zavidljiv rezultat. Kako deluje, podrobneje predstavimo v razdelku 3.5. Kasneje sta ista avtorja objavila še nadgrajeno različico [10], ki obljublja natančnost okoli dveh metrov. Glavna pomanjkljivost metode RADAR je, da je potrebno premeriti signale po celotnem območju, kar zahteva veliko časa. Z upoštevanjem sten³ [19] in modelom pojenja signala (angl. free space path loss model)⁴ je mogoče zgraditi model (angl. radio propagation model), s katerim nadomestimo fazo merjenja signalov, vendar so rezultati veliko slabši [10]. V zadnjih letih se poskuša z aktivnim učenjem in umetno inteligenco postopno izboljševati začetne modele [26] in tako zmanjšati število začetnih meritev signala. Kljub vsem izboljšavam, pa je še vedno potrebno pogosto meriti signale, saj majhna sprememba v prostoru močno vpliva na signale (npr. več ljudi, premik omare, nova dostopna točka, ipd.).

Vzporedno so se razvijali tudi algoritmi, ki delajo po principu trilateracije [21, 24, 27, 39], torej tako kot GPS. Glavna prednost trilateracije v primerjavi z RADAR-jem je v tem, da ni potrebno opraviti meritev signalov, ki zahtevajo veliko časa. Po drugi strani pa algoritem temelji na preračunavanju signala (RSSI [dBm]) v razdaljo [m], ki pa ni natančno. Najbolje se obnese model pojenja signala (angl. free-space path loss model), ki ga opišemo v razdelku 3.2.1. Ker razdalje med dostopno točko in izvorom signala ne znamo natančno izračunati, trilateracija ne vrača zadovoljivih rezultatov. Eno izmed implementacij trilateracije podrobno obrazložimo v razdelku 3.6.

Dr. Farid je s kolegi pripravil pregled vseh trenutnih metod za določanje lokacije v notranjih prostorih [17]. V članku so povzete vse metode, ki obljublajo najboljše rezultate.

3.2 Ocena razdalje in kvalitete signala

Standard IEEE 802.11 predpisuje, da mora vsaka brezžična kartica (angl. network interface card, NIC) sprejetemu Wi-Fi okvirju dodati glavo *radiotap*,

³Za iskanje števila sten med dostopno točko in specifično lokacijo, se uporablja *Cohen-Sutherland line-clipping* algoritem

⁴Podrobeneje je predstavljen v razdelku 3.2.1

ki vsebuje podrobnejše podatke o signalu (moč signala, nivo šuma, kvaliteta signala, vrednost TX, ipd.) [32]. V nadaljevanju se bomo osredotočili predvsem na naslednji dve polji:

SSI Signal [dBm] - v polje *radiotap.dbm_antsignal* se zapiše moč signala zaznana ob anteni brezžične kartice. Vrednost je predstavljena s predznačenim 8-bitnim številom.

SSI Noise [dBm] - v polje *radiotap.dbm_antnoise* se zapiše nivo šuma ob anteni brezžične kartice. Vrednost je predstavljena s predznačenim 8-bitnim številom.

Za prikaz moči signala se uporabljata dve različni enoti: mW (miliwatt) in dBm (decibel-miliwatt). Brezžične kartice, prisotne v današnjih računalnikih, oddajajo z močjo $100mW (= 20dBm)$ in zaznavajo signale vse do moči $2.5 \times 10^{-10}mW (= -96dBm)$. Moč izražena v mW hitro postane zelo majhna, saj moč pada s kvadratom razdalje, številke pa postanejo nepregledne. Zato se za signale pogosteje uporablja enota dBm. Za pretvorbo med enotami se uporablja naslednja formula:

Pretvorba iz moč $P[mW]$ v moč $P[dBm]$ in obratno:

$$P[dBm] = 10 * \log_{10}(P[mW]) \quad (3.1a)$$

$$P[mW] = 10^{\frac{P[dBm]}{10}} \quad (3.1b)$$

Vrednost SNR (angl. signal-to-noise ratio) upošteva tudi šum v okolici. Ta nam pove, ali je signal paketa dovolj dober, da lahko razumemo njegovo vsebino. Izračunamo ga lahko po naslednji formuli:

$$SNR[dBm] = P_{signal}[dBm] - P_{noise}[dBm] \quad (3.2)$$

Nivo šuma (angl. signal noise level) se ponavadi giblje med -95 in -92 dBm. Da se podatki lahko prenašajo med oddajnikom in sprejemnikom, mora biti SNR vsaj 10 dBm. V nasprotnem primeru je signal prešibek, saj se ga težko razloči od šuma.

0000	00 00 19 00 6f 08 00 00	02 34 19 73 00 00 00 00	...o... .4.s...
0010	10 02 99 09 80 04 37 a1	00 40 00 00 00 ff ff ff	...@.....
0020	ff ff ff 54 ae 27 69 58	41 ff ff ff ff ff ff 90	...T.'iX A.....
0030	d4 00 0b 53 74 65 61 6d	52 6f 6c 6c 65 72 01 04	...Steam Roller..
0040	02 04 0b 16 32 08 0c 12	18 24 30 48 60 6c 03 012... .50H'l..
0050	09 2d 1a 20 00 1a ff 00	00 00 00 00 00 00 00 00
0060	00 00 00 00 00 00 00 00	00 00 00 00 00 7f 04 00
0070	00 08 04 6b 07 0f ff ff	ff ff ff ff dd 09 00 10	...k...
0080	18 02 01 00 00 00 00 dd	1e 00 90 4c 33 20 00 1aL3 ..
0090	ff 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00
00a0	00 00 00 00 00 00 dd	07 00 50 f2 08 00 11 00P.....
00b0	d7 e6 61 63		..ac

Slika 3.1: Probe request okvir z označeno glavo *radiotap*. Bajt označen z rdečo barvo predstavlja *SSI Signal*, z zeleno pa *SSI Noise*.

S programom Wireshark lahko preberemo vse okvirje, ki jih zazna brezžična kartica. Na sliki 3.1 vidimo primer zajetega okvirja. Gre za *probe request*, poslan z mobilnega telefona, s katerim telefon poskuša ugotoviti, ali je v bližini omrežje s SSID-jem *SteamRoller*. S svetlo modro barvo je označena glava *radiotap*, ki jo je dodala brezžična kartica. Ta med drugim vsebuje tudi podatka o kvaliteti signala (rdeč okvirček) in šumu (zelen okvirček) v okolici. Obe številki sta predstavljeni s predznačenim 8-bitnim številom, ki ju lahko pretvorimo v desetiški sistem po naslednjem postopku⁵:

$$d7_{hex} = 11010111_2 = 215 \rightarrow 215 - 256 = -41 \quad (3.3a)$$

$$a1_{hex} = 10100001_2 = 161 \rightarrow 161 - 256 = -95 \quad (3.3b)$$

Brezžična kartica je torej zaznala moč signala -41 dBm ob šumu -95 dBm. Izračunamo lahko še SNR (signal-to-noise ratio):

$$SNR = P_{signal} - P_{noise} = -41dBm - (-95dBm) = 54dBm \quad (3.4)$$

Vrednost SNR znaša 54 dBm, kar pomeni, da je signal zelo dober in primeren za normalno komunikacijo med sprejemnikom in oddajnikom. Rezultati niso presenetljivi, saj je bila razdalja med mobilnim telefonom in sprejemnikom manj kot 1 meter.

⁵Ker gre za predznačeni števili, je potrebno odšteti 256.

V znanstvenih člankih [9, 24, 33] opozarjajo na nihanje moči signala. Na izmerjeno moč namreč vplivajo številni dejavniki: ovire na poti, obrnjenost naprave, ki oddaja signal, natančnost senzorja, prisotnost šuma, ipd. Izmerjena moč signala lahko v zelo kratkih časovnih intervalih odstopa tudi za 5 dBm in več. Ker naprave ponavadi v kratkem časovnem intervalu pošljejo več paketov, lahko podatke povprečimo in zmanjšamo vpliv nihanja. V člankih [9, 33] predlagajo izračun mediane, saj s tem zmanjšamo vpliv vrednosti, ki močno odstopajo od povprečja. Primer zaznanih moči signala, ki jih je oddal mobilni telefon v nekaj milisekundah:

$$-40, -39, -39, -39, -38, \mathbf{-57}, \mathbf{-56}, -38, -39, -38 \quad [dBm] \quad (3.5)$$

Aritmetična sredina: -42.3 dBm

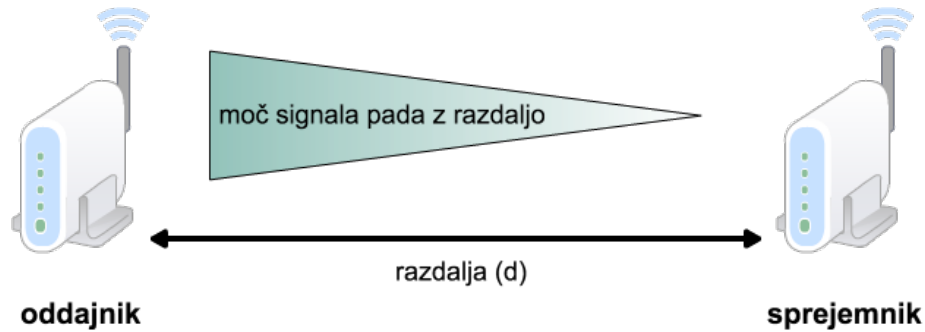
Mediana: -39 dBm

Aritmetična sredina močno odstopa od večine ostalih meritev. Vmes sta namreč dve meritvi (v krepkem tisku), ki močno vplivata na izračun povprečja. Na mediano pa skrajne vrednosti nimajo velikega vpliva. Ker je narava problema bližje obnašanju mediane, se uporaba mediane v praksi tudi boljše obnese.

V nadaljevanju bomo predstavili, kako lahko na podlagi moči signala ocenimo razdaljo. Primerni sta dve vrednosti: moč signala in SNR. V praksi se boljše izkaže moč signala, saj je SNR preveč odvisen od šuma v okolici. To potrjujejo tudi izsledki raziskave, objavljeni v članku [9].

3.2.1 Model pojenja signala

Oddajnik Wi-Fi oddaja signale v obliki elektromagnetnega valovanja. Moč signala je najvišja tik ob anteni, nato pa pojenja z razdaljo. Kako razdalja vpliva na signal, opisuje model pojenja signala z razdaljo skozi prazen prostor (angl. free-space path loss model), ki predpostavlja, da je vidna linija (angl. line of sight) med oddajnikom in sprejemnikom brez ovir. Vsaka ovira



Slika 3.2: Oddajnik in sprejemnik: pojenjanje signala z razdaljo.

na poti, lahko še dodatno odbije oz. razprši elektromagnetno valovanje, kar vpliva na moč signala. Izguba signala (angl. path loss) je premo sorazmerna s kvadratom razdalje med oddajnikom in sprejemnikom in s kvadratom frekvence signala, ki ga oddaja oddajnik. Fizikalna definicija za izgubo signala je:

$$path_loss = \left(\frac{4\pi d^{\frac{n}{2}}}{\lambda}\right)^2 = \left(\frac{4\pi f d^{\frac{n}{2}}}{c}\right)^2 \quad (3.6)$$

d razdalja med oddajnikom in sprejemnikom,

lambda valovna dolžina, $\lambda = \frac{c}{f}$,

n eksponent izgube signala, $n = 2$ predstavlja prazen prostor (zrak),

c svetlobna hitrost, $299792458 \frac{m}{s}$,

f frekvenca signala.

Da bomo lažje preoblikovali formulo v željeno obliko, jo logaritmirajmo po formuli 3.1a:

$$\log(path_loss) = 20\log\left(\frac{4\pi}{c}\right) + 20\log(f) + 10n\log(d) \quad (3.7)$$

Konstanto lahko izračunamo $20\log\left(\frac{4\pi}{c}\right) = -27,55$, enačbo pa poenostavimo:

$$\log(path_loss) = 20\log(f) + 10n\log(d) - 27,55 \quad (3.8)$$

Moč signala pada z razdaljo. Največja je pri oddajniku, označimo jo z $P_o[dBm]$. Z upoštevanjem formule za pojenjanja signala lahko izračunamo moč pri sprejemniku $P_s[dBm]$:

$$P_s = P_o - \log(path_loss) = P_o - 20\log(f) - 10n\log(d) + 27,55 \quad (3.9)$$

Iz te zveze lahko izrazimo razdaljo $d[m]$:

$$d = 10^{\frac{P_o - P_s + 20\log(f) + 27,55}{10n}} \quad (3.10)$$

Iz formule je razvidno, da je razdalja odvisna od štirih spremenljivk. P_s predstavlja moč, preostale tri pa lahko fiksiramo. Poglejmo njihove privzete vrednosti:

P_o - moč signala pri oddajniku [dBm]. Večina brezžičnih kartic oddaja z močjo 100 mW, kar znaša **20 dBm**;

f - frekvenca valovanja [MHz]. Ta je zapisana znotraj glave *radiotap* v polju *Channel frequency*, poleg *SSI Signal* in *SSI Noise*. WLAN uporablja 13 različnih kanalov⁶. Prvi kanal uporablja frekvenco 2412 MHz, trinajsti pa 2472 MHz. Za privzeto vrednost smo zato izbrali srednjo vrednost, t.j. **2432 MHz**;

n - eksponent izgube signala. Za n izberemo 2, kadar predvidevamo, da se bo signal širil po prostoru brez ovir in bo prispel do sprejemnika po vidni liniji (angl. line-of-sight). Izbira eksponenta je odvisna od števila ovir, ki se nahajajo med oddajnikom in sprejemnikom. Močan vpliv na izgubo signala imajo stene (debelina in material). V članku [33] so s

⁶Število kanalov je odvisno glede na zakonodajo. V Evropi se uporablja 13 kanalov, v Severni Ameriki pa zgolj 11.

praktičnim primerom pokazali, kateri eksponent je v zaprtih prostorih najprimernejši. Predlagajo $n = 5$. Na sliki 3.3 sta prikazani krivulji za $n = 2$ in $n = 5$, ki ponazarjata padec moči signala z razdaljo.

Algoritem za določanje razdalje med oddajnikom in sprejemnikom smo implementirali v programskem jeziku Python. Upoštevali smo tudi privzete vrednosti za 3 parametre, ki pa se lahko poljubno spreminjajo:

```

1 def convert_rssi_to_distance(Pr, Pt=20, n=5, f=2432):
2     """
3     Pr = power at receiver [dBm]
4     Pt = power at transmitter [dBm]
5     n = path loss exponent, n = 2 represents free space
6     f = frequency [MHz]
7     Return: distance between transmitter and receiver [meters]
8     """
9     return 10**((Pt - Pr - 20 * log(f, 10) + 27.55) / (10 * n))

```

3.2.2 Kvaliteta signala

Operacijski sistemi moči signala uporabniku navadno ne pokažejo v mW ali dBm, temveč jo vsak na svoj način preračuna v kvaliteto signala, ki jo širša množica ljudi boljše razume. Obstajajo različne definicije, kako prevesti moč signala v kvaliteto signala, najbolj razširjena je definicija, ki jo med drugim uporablja tudi Microsoft v operacijskem sistemu Windows [2].

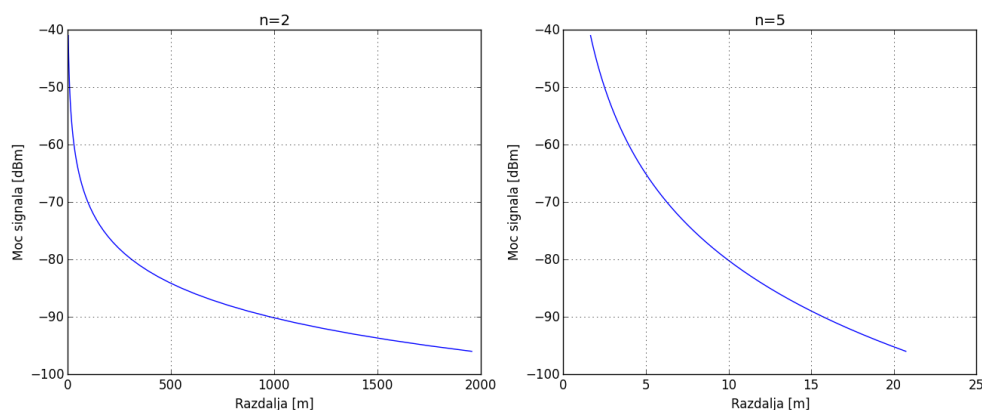
Kvaliteta signala - vrednost izražena v odstotkih. Vrednost 0 % predstavlja moč signala -100 dBm, 100 % pa -50 dBm. Vmesne vrednosti se izračunajo z linearno interpolacijo.

Kvaliteto signala Q lahko izračunamo s pomočjo naslednje formule:

$$Q[\%] = 2 \times (P[dBm] + 100), P[dBm] \in [-100, -50] \quad (3.11)$$

Izračun moči signala iz kvalitete signala:

$$P[dBm] = \frac{Q}{2} - 100, Q \in [0, 100] \quad (3.12)$$



Slika 3.3: Primerjava upadanja moči signala z razdaljo pri $n = 2$ in $n = 5$.

V programskem jeziku *Python* lahko pretvorbo zapišemo takole:

```

1 def convert_rssi_to_quality(rssi_value):
2     if rssi_value <= -100:
3         quality = 0
4     elif rssi_value >= -50:
5         quality = 100
6     else:
7         quality = 2 * (rssi_value + 100)
8     return quality

```

3.2.3 Primeri izračunov

Pripravili smo 7 različnih vrednosti moči signala. Za vsako smo ocenili razdaljo in kvaliteto. Vključili smo tudi skrajne vrednosti, z namenom, da preverimo robne pogoje.

Rezultati so predstavljeni v tabeli 3.1. V prvem stolpcu je izražena moč signala v dBm, v drugem pa v mW. Slednja števila so zelo majhna in jih v bodoče ne bomo več uporabljali. V tretjem in četrtem stolpcu sledijo razdalje, ki smo jih preračunali po dveh različnih modelih pojenjanja signala. Na sliki 3.3 je prikazana primerjava upadanja moči signala z razdaljo pri

moč [dBm]	moč [mW]	razdalja ⁷ [m]	razdalja ⁸ [m]	kvaliteta [%]
-41	7.9×10^{-5}	11	2.6	100
-50	1.0×10^{-5}	31	3.9	100
-60	1.0×10^{-6}	98	6.3	80
-75	3.2×10^{-8}	549	12.4	50
-80	1.0×10^{-8}	977	15.7	40
-96	2.5×10^{-10}	6162	32.8	8
-100	1.0×10^{-10}	9760	39.4	0

Tabela 3.1: V tabeli so prikazane ocene razdalje in kvalitete signala za 7 različnih moči.

$n = 2$ in $n = 5$.

Odprt prostor ($n = 2$): ta model predpostavlja, da med sprejemnikom in oddajnikom ni ovir. Moč signala z razdaljo pada zelo počasi. Majhna sprememba v moči signala pa lahko pomeni spremembo v razdalji tudi za več kot 100 m;

Zaprta prostor ($n = 5$): ta model predpostavlja, da so med sprejemnikom in oddajnikom ovire, ki odbijajo elektromagnetno valovanje. Posledično moč signala z razdaljo hitreje pada.

Zadnji stolpec predstavlja kvaliteto signala, izračunano po formuli 3.11.

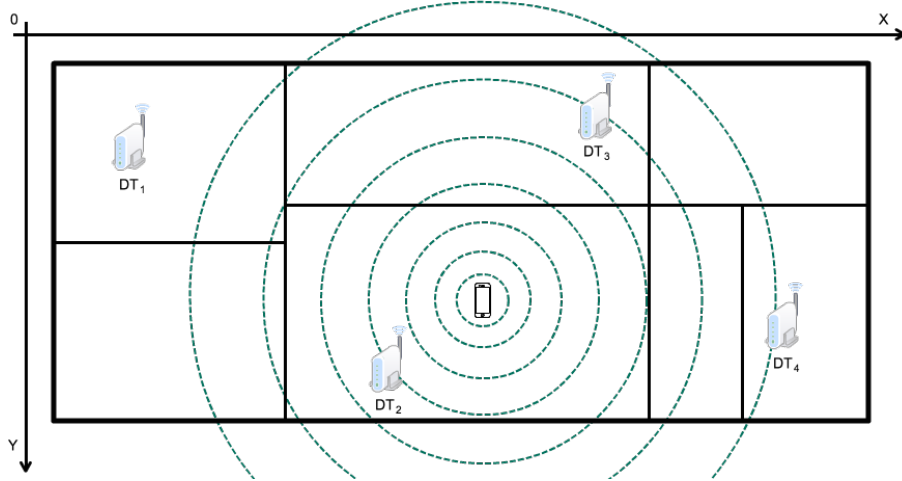
3.3 Problem določanja lokacije

V naslednjih razdelkih bomo opisali tri načine določanja lokacije v zaprtih prostorih. Za lažjo predstavo smo pripravili generičen prostor, razdeljen na več manjših prostorov. Skica se nahaja na sliki 3.4. V enem izmed prostorov se nahaja mobilna naprava, ki oddaja signale Wi-Fi. V okolici se nahajajo štiri dostopne točke (DT_1, DT_2, DT_3, DT_4), za katere poznamo lokacijo.

⁷Razdalja v odprtem prostoru ($n=2$).

⁸Razdalja v zaprtem prostoru ($n=5$).

Prostor smo postavili v 2D kartezični koordinatni sistem, z izhodiščem v zgornjem levem kotu. Lokacije dostopnih točk lahko torej opišemo s koordinatami: $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)\}$.



Slika 3.4: Problem določanja lokacije.

Zanima nas lokacija l mobilne naprave, ki oddaja signale Wi-Fi. Predstavljena naj bo s koordinatama: $l = (x, y)$.

V nadaljevanju bomo opisali tri metode, ki se pogosto uporabljajo za reševanje opisanega problema.

3.4 Metoda najmočnejše dostopne točke

Metoda najmočnejše dostopne točke (angl. strongest base station)⁹ velja za najenostavnejšo metodo za določanje lokacije [22]. Izmed vseh dostopnih točk, ki zaznajo signal, vrne lokacijo tiste, ki je zaznala najmočnejši signal [35]. Ta metoda je sicer enostavna, vendar vrača le lokacije dostopnih točk. Ločljivost je torej zelo majhna. Metodo uporabimo, če imamo na voljo veliko dostopnih točk. To pa po drugi strani pomeni, da je tak sistem za določanje lokacije zelo drag, večkrat pa zaradi prevelike zasedenosti prostora tudi nepraktičen.

⁹V literaturi se pogosto pojavlja tudi pod imenom metoda najbližje dostopne točke (angl. nearest base station).

3.5 Profiliranje signalov Wi-Fi

Leta 2000 sta dr. Bahl in dr. Padmanabhan, iz raziskovalnega oddelka podjetja Microsoft, objavila članek [9], v katerem sta opisala RADAR. Gre za prvo metodo za določanje lokacije v notranjih prostorih, ki temelji na moči signala Wi-Fi. Še danes velja za eno najnatančnejših metod za določanje lokacije, članek pa citirajo številni znanstveniki po vsem svetu. V nadaljevanju bomo predstavili glavno idejo metode.

Da bo opis enostavnejši, se bomo osredotočili na sistem s tremi dostopnimi točkami. Te je potrebno razporediti po prostoru tako, da celotno območje prekrije signal vseh treh točk. Nato uvedemo koordinatni sistem, v katerega umestimo 2D tloris prostorov. RADAR sestavljata dve fazi:

(1) faza učenja (angl. offline learning phase). V tej fazi se opravijo meritve signalov na različnih lokacijah. Vsako lokacijo se določi s koordinato (x, y) in terko signalov $(S_1, S_2, S_3)^{10}$, s katero opišemo izmerjene signale do vseh treh dostopnih točk. Meritve je treba opraviti večkrat in jih nato povprečiti. Po koncu meritev imamo na voljo bazo lokacij s pripadajočimi meritvami signalov. Ta proces smo poimenovali profiliranje signalov Wi-Fi (angl. finger-printing).

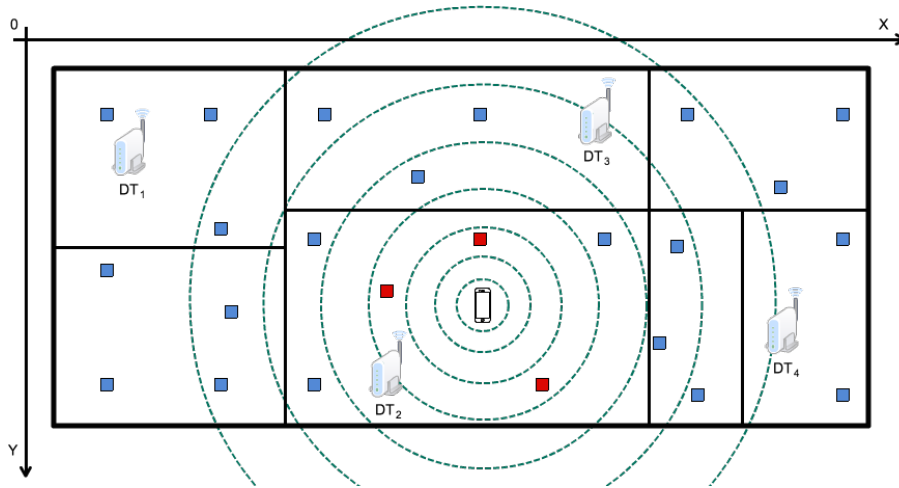
(2) faza določanja lokacije (angl. online estimation phase). Za prejet signal $s' = (S'_1, S'_2, S'_3)$ se s pomočjo opravljenih meritev izračuna okvirna lokacija zaznanega signala. Avtorja sta definirala algoritem *Nearest neighbours in signal space (NNSS)*, ki za sprejet signal s' poišče najbližjo meritev s s pomočjo evklidske razdalje:

$$d(s', s) = \sqrt{(S'_1 - S_1)^2 + (S'_2 - S_2)^2 + (S'_3 - S_3)^2} \quad (3.13)$$

Z uporabo te metode pridobimo spodbudne rezultate, katere pa lahko hitro in enostavno izboljšamo tako, da pri izračunu poiščemo k naj-

¹⁰S = signal

bližjih meritev, ki jih nato povprečimo. Avtorja predlagata $k = 3$, kar rezultate izboljša do 20 %.



Slika 3.5: Določanje lokacije s profiliranjem signalov Wi-Fi. Kvadratici predstavljajo lokacije meritev. Z rdečimi kvadratici smo označili 3 meritve, ki so glede na evklidsko razdaljo najmanj oddaljene od signalov, ki jih je oddala mobilna naprava. Rezultat 2. faze je centroid treh rdečih kvadratkov.

Pravzaprav je algoritem kompleksnejši. Na vsaki lokaciji predvideva merjenje signala za štiri različne usmeritve mobilne naprave (sever, jug, vzhod, zahod). Izkaže se, da če se človek nahaja med dostopno točko in napravo, se lahko meritve razlikujejo tudi za $3,5\text{dBm}$.

Kasneje sta avtorja v novem članku [10] objavila izboljšano verzijo algoritma, ki obljublja 30 % izboljšave. V članku uvedeta več modelov, ki jih ločita glede na število ljudi v prostoru. Razložita tudi, koliko dostopnih točk je potrebnih za optimalno delovanje metode.

Čeprav RADAR obljublja natančnost do nekaj metrov (2-3 m), je glavna slabost še vedno prevelika količina potrebnega časa za meritve signalov. V primeru, da se premakne ena izmed dostopnih točk, so potrebne nove meritve, ki zopet zahtevajo svoj čas. Da bi odpravili to omejitev, lahko za gradnjo modela uporabimo matematične postopke, ki upoštevajo razpored

sten in štejejo število prostorov med dostopno točko in lokacijo s pomočjo *Cohen-Sutherland line-clipping* algoritma. Ti matematični modeli resda poenostavijo prvo fazo, vendar pa je natančnost skoraj dvakrat slabša [9].

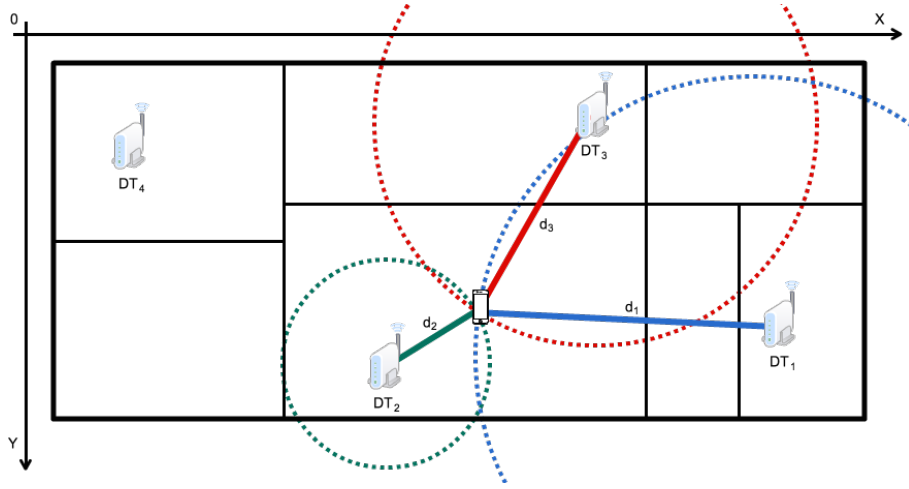
3.6 Trilateracija

Trilateracija je matematični postopek, ki na podlagi treh znanih točk in treh izmerjenih razdaljah, omogoča izračun neznane lokacije. Beseda *tri* na začetku predstavlja 3 znane točke. Posplošena trilateracija, ki omogoča izračun na podlagi N znanih točk, se imenuje multilateracija. Pogosto pa se v literaturi, npr. [15], z besedo trilateracija označuje tudi posplošen primer, zato bomo tudi mi v nadaljevanju ves čas govorili o trilateraciji.

Trilateracija se uporablja takrat, ko znamo natančno izmeriti razdalje med znanimi lokacijami in neznano lokacijo. Pri GPS je možno na podlagi zakasnitve signala med satelitom in napravo, izračunati razdaljo po formuli $d = c * (t_2 - t_1)$, kjer c predstavlja svetlobno hitrost s katero potujejo signali, t_2 čas sprejema signala, t_1 pa čas oddaje. Za določitev 3D lokacije so potrebni štirje sateliti, za 2D pa trije. V splošnem torej potrebujemo $n + 1$ znanih točk, da lahko določimo lokacijo v n -dimenzionalnem prostoru.

Na podoben način pa lahko lokacijo določimo z Wi-Fi. Namesto o satelitih govorimo o dostopnih točkah. Na podlagi moči signala, lahko s pomočjo modela za pojenjanje signala, ki smo ga podrobneje predstavili v razdelku 3.2.1, izračunamo okvirno razdaljo med dostopno točko in izvorom signala. Nato pa z matematičnim postopkom določimo neznano lokacijo. Obstaja več različnih načinov, nekateri izmed njih so opisani v naslednjih znanstvenih člankih [24, 27, 39]. V nadaljevanju na kratko povzamemo način, ki je podrobneje predstavljen v članku [21].

Da bo postopek lažje razumljiv, se osredotočimo na sistem s tremi dostopnimi točkami DT_1 , DT_2 in DT_3 , ki se nahajajo na lokacijah (x_1, y_1) , (x_2, y_2) in (x_3, y_3) . Poznamo tudi razdalje d_1 , d_2 in d_3 med dostopnimi točkami in neznano lokacijo (x, y) . Ker poznamo razdalje za vsako dostopno točko do



Slika 3.6: Na podlagi treh znanih lokacij (DT_1 , DT_2 in DT_3) in treh razdalj (d_1 , d_2 in d_3) med dostopnimi točkami in iskano lokacijo, lahko s postopkom trilateracije določimo lokacijo izvora signala.

izvora signala, lahko za vsako zapišemo enačbo, ki opisuje sfero, na kateri se nahaja neznana lokacija. Presečišče sfer oz. približek predstavlja iskano lokacijo.

$$\begin{aligned}(x - x_1)^2 + (y - y_1)^2 &= d_1^2 \\(x - x_2)^2 + (y - y_2)^2 &= d_2^2 \\(x - x_3)^2 + (y - y_3)^2 &= d_3^2\end{aligned}$$

Ta sistem je možno rešiti na več načinov. Npr. avtorji članka [21] predlagajo naslednjo različico. Lokacijo (x, y) lahko izračunamo z naslednjima formulama:

$$\begin{aligned}x &= \frac{1}{\Delta}(2T_1(y_1 - y_3) - 2T_2(y_1 - y_2)) \\y &= \frac{1}{\Delta}(2T_2(x_1 - x_2) - 2T_1(x_1 - x_3))\end{aligned}$$

kjer so prisotni parametri Δ , T_1 in T_2 , te pa lahko izračunamo z naslednjimi formulami:

$$\Delta = 4((x_1 - x_2)(y_1 - y_3) - (x_1 - x_3)(y_1 - y_2))$$

$$T_1 = d_2^2 - d_1^2 - x_2^2 + x_1^2 - y_2^2 + y_1^2$$

$$T_2 = d_3^2 - d_1^2 - x_3^2 + x_1^2 - y_3^2 + y_1^2$$

Enačba je rešljiva, ko $\Delta \neq 0$, torej vedno, ko dostopne točke niso vse na isti premici. Ta primer izračuna je le eden izmed mnogih. Pogosto se uporablja tudi metoda najmanjših kvadratov (angl. least square).

Natančnost trilateracije ni tako dobra kot pri GPS, saj signal v notranjih prostorih motijo številne ovire, kar posledično vpliva na preračun razdalj. Kljub temu pa ima eno veliko prednost pred profiliranjem signalov Wi-Fi. Pri vzpostavitvi sistema namreč niso potrebe meritve signalov, kar poceni in pohitri proces. Opravka pa imamo z vzpostavitvijo koordinatnega sistema in določanjem lokacij dostopnih točk.

3.7 Primerjava

Sledi kratek povzetek in primerjava metod, ki smo jih predstavili v prejšnjih treh razdelkih: *metoda najmočnejše dostopne točke* (MNDT), *profiliranje signalov Wi-Fi* (PSW) in *trilateracija*. Primerjali jih bomo glede na naslednje tri parametre:

- **ločljivost** - ali metode vračajo koordinate (x, y) ali le lokacije dostopnih točk;
- **cena** - koliko je dela z vzpostavitvijo in kalibracijo sistema ter kakšna je posledično cena;
- **natančnost** - okvirna natančnost, ki jo obljublja različne implementacije predstavljene v znanstvenih člankih.

Metoda najmočnejše dostopne točke je najosnovnejša metoda in predstavlja referenčni mejnik (angl. baseline) za vse ostale metode. Glavna pomanjkljivost je slaba ločljivost, ker vrača le lokacije dostopnih točk. Po drugi strani pa je vzpostavitev enostavna, saj ni začetne kalibracije, prav tako pa ni potrebe po vpeljavi koordinatnega sistema. Natančnost je v splošnem slaba. Lahko pa je tudi dobra, če imamo na voljo veliko dostopnih točk, vendar pa v tem primeru naraste cena, saj je potrebno namestiti veliko dostopnih točk.

Profiliranje signalov Wi-Fi je metoda, ki ima dobro ločljivost, saj vrača koordinate (x, y) in ne zgolj lokacije dostopnih točk. Glavno pomanjkljivost predstavlja dolgotrajna kalibracija sistema, saj je v prvi fazi potrebno veliko časa za merjenje moči signalov. V praksi ponuja tudi najboljšo natančnost izmed vseh metod. Npr. sistem RADAR ob pravilni postavitvi dostopnih točk in kvalitetni izmeri signalov, omogoča izračun lokacije do dveh metrov natančno. Vendar pa da lahko sistem ves čas zagotavlja tako visoko natančnost, je potrebno večkrat opraviti meritve signalov, predvsem takrat, ko pride do kakšne spremembe v prostoru, npr. do premika omare ali pa montaže nove dostopne točke.

Trilateracija prav tako ponuja dobro ločljivost. Glavna prednost predstavlja enostavnost vzpostavitve, saj ni potrebe po kalibraciji in meritvah signalov. Potrebno je le vpeljati koordinatni sistem in nanj vnesti lokacije dostopnih točk. Vendar pa za ceno enostavnosti, peša natančnost. Glede na izsledke iz številnih člankov, je realna natančnost nekje med 5-10 m, odvisno od razporeditve prostorov, ovir in števila dostopnih točk.

Skratka, vsaka metoda ima svoje prednosti in slabosti. Podrobnejši pregled vseh metod za določanje lokacij v notranjih prostorih je predstavljen v [17]. Poleg tehnologije Wi-Fi so avtorji pregledali tudi ostale tehnologije (Bluetooth, RFID, ZigBee, FM) in zaključujejo, da je Wi-Fi trenutno glede na ceno najprimernejša tehnologija za določanje lokacije.

metoda	ločljivost	cena	natančnost
MNDT ¹¹	-	+	- ¹²
PSW ¹³	+	-	+ (2-3 m)
Trilateracija	+	o	o (5-10 m)

Tabela 3.2: Kratek povzetek prednosti oz. pomanjkljivosti metod za določanje lokacije v notranjih prostorih z Wi-Fi.

V nadaljevanju magistrske naloge se bomo posvetili razvoju nove metode, ki bo združevala prednosti vseh opisanih metod z manjšimi kompromisi. Nova metoda mora omogočati enostavno vzpostavitev, solidno natančnost in ločljivost vsaj na nivoju regij oz. prostorov.

¹¹Metoda najmočnejše dostopne točke

¹²Če imamo na voljo veliko dostopnih točk, potem je natančnost lahko dobra. To pa pomeni, da je tak sistem drag. V splošnem natančnost ni dobra.

¹³Profiliranje signalov Wi-Fi

Poglavje 4

Pregled produktov za določanje lokacije

Številni proizvajalci brezžične opreme ponujajo rešitve, ki omogočajo določanje lokacije v notranjih prostorih s pomočjo tehnologije Wi-Fi. V nadaljevanju podrobneje predstavimo tri uveljavljene produkte: Cisco MSE, Meraki Presence Analytics in Ruckus SPoT. Ob koncu poglavja vse tri produkte primerjamo med seboj glede na: ceno, programski vmesnik, zahtevnost kalibracije, kvaliteto analitike, natančnost določanja lokacije, enostavnost uporabe ter preverimo, ali se rešitev nahaja v oblaku ali lokalno.

4.1 Cisco MSE

Podjetje Cisco Systems velja za največje podjetje na področju omrežne tehnologije. Ponujajo različne rešitve za določanje lokacije v notranjih prostorih in te oglašujejo pod imenom *Connected Mobile Experiences* (CMX). Med njimi je tudi produkt MSE (Mobility Services Engine), ki omogoča določanje lokacije s tehnologijo Wi-Fi.

Osnovno opremo predstavljajo dostopne točke Cisco. Te se nato preko ostale infrastrukture povežejo z napravo MSE, ki je lahko fizična oz. virtualna. Če uporabljamo virtualno različico, potem je potrebno zagotoviti kar

POGLAVJE 4. PREGLED PRODUKTOV ZA DOLOČANJE LOKACIJE

16 GB delovnega pomnilnika. MSE zmore spremljati do 5000 naprav hkrati. Preden lahko sistem določi lokacijo naprav, ga je potrebno še ustrezno nastaviti. V sistem je potrebno vnesti zemljevid notranjih prostorov (angl. floorplan), na katerem se označi lokacije dostopnih točk. Poleg dostopnih točk, se lahko definira tudi regije, stene, vrata, itd. Več podatkov kot vnesemo, boljša je natančnost. Sistem samo na podlagi teh nastavitev omogoča zaznavo lokacij mobilnih naprav. Vendar pa, če želimo doseči sprejemljivo natančnost, je potrebno opraviti natančno kalibracijo, t.i. profiliranje signalov Wi-Fi (angl. finger-printing). Cisco v ta namen ponuja programsko opremo, ki se jo namesti na prenosni računalnik. Program nato pokaže na lokacije, kjer je potrebno izmeriti signal. Ko se postavimo na določeno lokacijo, pritisnemo gumb start. Med tem časom prenosni računalnik generira signale Wi-Fi, sistem pa opravlja meritve. Za vsako meritev potrebujemo 2-3 minute časa. Na ta način se zgradi karta signalov (angl. radio map), ki omogoča boljšo natančnost. Cisco ne izdaja informacije o tem, na kakšen način se določi lokacija. Glede na to, da pa zahteva kalibracijo s profiliranjem signalov Wi-Fi, sklepamo, da uporablja algoritem, podoben RADAR-ju, ki smo ga podrobno opisali v razdelku 3.5.

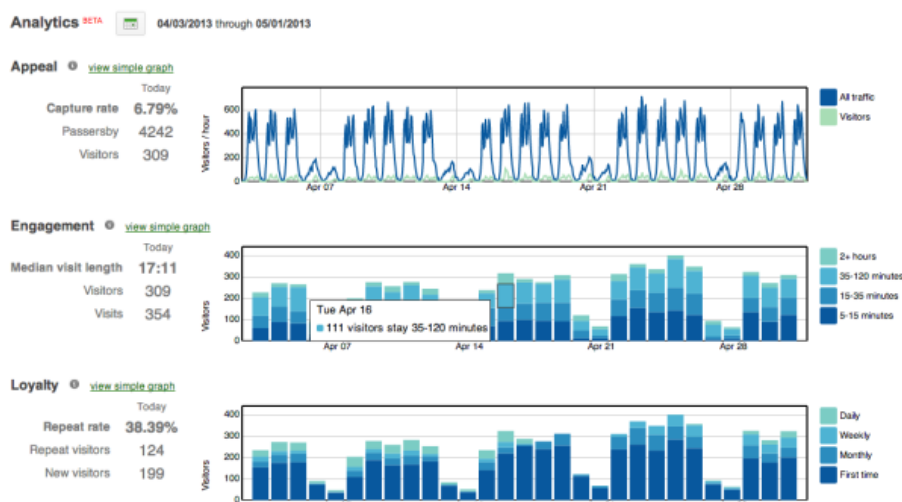
MSE poleg določanja lokacije, ponuja tudi grafični vmesnik, ki omogoča pregled osnovnih statistik. Gre za neke vrste analitični modul, ki ponuja natančnejši vpogled, kaj se dogaja v okolici dostopnih točk. Pregledujemo lahko toplotne karte (angl. heatmap), obisk skozi čas, kje se naprave pogosto zadržujejo, kako pogosto se naprave vračajo itd.

Poleg grafičnega vmesnika, je do podatkov možno dostopati tudi preko programskega vmesnika REST. Ta omogoča pridobiti tudi zemljevide prostorov in koordinate trenutnih lokacij za vsako napravo. Vmesnik je namenjen predvsem razvijalcem novih aplikacij, s katerimi je možno rešiti številne poslovne probleme.

Da lahko vzpostavimo tak sistem, je potrebno kupiti fizično opremo - dostopne točke, usmerjevalnike, stikala in MSE. Poleg pa je potrebno kupiti še ustrezno licenco, cena pa je določena s številom spremljanih naprav.

4.2 Meraki Presence Analytics

Podjetje Meraki je bilo ustanovljeno leta 2006 na inštitutu MIT. Razvijajo strojno opremo za žična in brezžična omrežja. Podjetje odlikujejo zelo enostavni produkti, za katere ni potrebe po vrhunskem znanju s področja računalniških omrežij. Ko se registrira nova Meraki oprema, se ta poveže na splet, tam pa se kreira oblachna storitev, ki poskuša čim več stvari nastaviti avtomatsko. Primarno so rešitve, ki jih ponuja Meraki, namenjene malim in srednje velikim podjetjem.

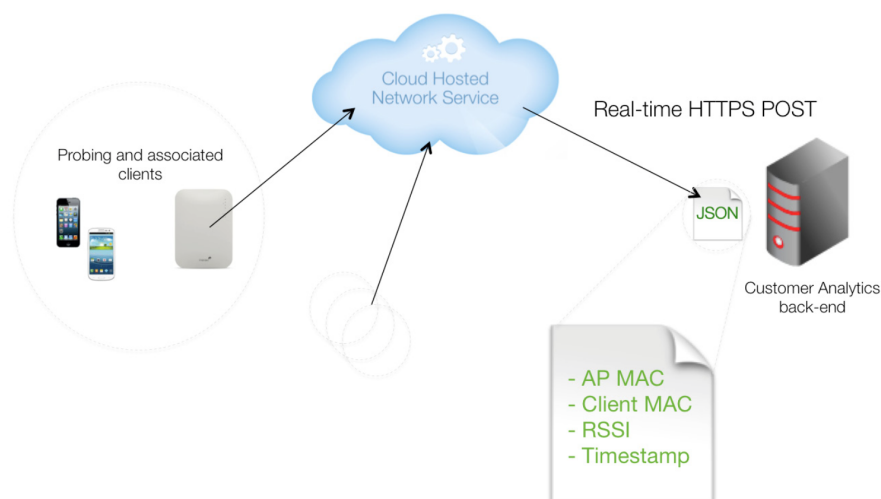


Slika 4.1: Analiza števila obiskovalcev, povprečnega časa obiska in kako pogosto se vračajo obiskovalci. Vir: Meraki White Paper [13]

Poleg klasičnih omrežnih storitev, pa Meraki ponuja tudi storitev, ki omogoča spremljanje dogajanja v bližini dostopnih točk. Imenuje se *Presence Analytics*. Vse dostopne točke v omrežju, ves čas zajemajo signale Wi-Fi, ki jih nato pošiljajo v oblak. Tam se podatki agregirajo iz vseh dostopnih točk, nato pa jih analitični modul obdelava ter generira grafe. Ti prikazujejo obisk skozi čas, povprečen čas zadrževanja (angl. engagement analytics), toplotne karte (angl. heatmap) in kako pogosto se naprave vračajo (angl. loyalty analytics) - na sliki 4.1 so prikazani primeri omenjenih grafov. Ta produkt je torej bolj kot določanju lokacije, namenjen analiziranju dogajanja v okolici

POGLAVJE 4. PREGLED PRODUKTOV ZA DOLOČANJE LOKACIJE

dostopnih točk. Meraki je pripravil tudi vmesnik REST, ki omogoča dostop do podatkov o moči signala, ki so ga zaznale dostopne točke. Na sliki 4.2 je predstavljena arhitektura te rešitve.



Slika 4.2: Arhitektura za *Presence Analytics*. Vir: Meraki White Paper [13]

Pred kratkim je Meraki predstavil novo storitev, ki omogoča tudi določanje lokacije. Vmesnik REST po novem vrača tudi (x,y) koordinate. Še vedno pa ni možno vnesti kalibracijske mape. Posledično je natančnost določanja lokacije še vedno zelo omejena.

Meraki ves čas poudarja tudi zasebnost. Javno so objavili rešitev [13], ki omogoča skrivanje naslovov MAC. S pomočjo zgoščevalne funkcije *SHA1* zgostijo vsak naslov MAC, skupaj z identifikatorjem organizacije. Na koncu shranijo le štiri desne bajte:

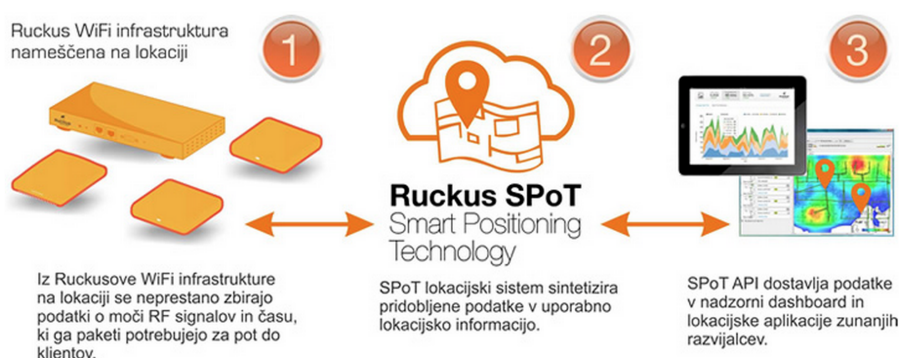
$$SHA1(naslov\ MAC \parallel sol\ organizacije).desniStirjeBajti() \quad (4.1)$$

Meraki ponuja cenovno ugodnejšo rešitev od Cisca. Sicer gre za precej različna produkta, vendar pa z novo verzijo vmesnika REST, tudi Meraki omogoča lociranje naprav Wi-Fi. Meraki pogosto naleti na kritike, saj uporaba njihovih produktov ni možna brez pošiljanja podatkov v oblak. Rešitve

tako za nekatera podjetja, ki imajo stroge politike glede podatkov, ne pridejo v poštev. Kljub temu pa podjetje zelo hitro raste, predvsem na račun svoje enostavne kalibracije. Podjetje je v celoti odkupila korporacija Cisco Systems za 1,2 milijardi dolarjev. Po novem se podjetje imenuje Cisco Meraki.

4.3 Ruckus SPoT

Podjetje Ruckus Wireless razvija tehnologijo na področju brezžičnih omrežij. Njihovi produkti se pogosto uporabljajo v velikih podjetjih. Pred kratkim so začeli s prodajo nove storitve, ki se imenuje *SPoT* (Smart Positioning Technology). Gre za oblačno storitev, ki omogoča določanje lokacije v notranjih prostorih, na podlagi signala Wi-Fi.



Slika 4.3: Shema sistema Ruckus SPoT. Vir: telos.si

Na sliki 4.3 je natančneje predstavljena storitev *SPoT*. Dostopne točke¹ pošiljajo podatke v oblak, kjer se agregirajo. Preden storitev lahko izračuna lokacije mobilnih naprav, je potrebno sistem še ustrezno nastaviti, podobno kot pri storitvi, ki jo ponuja Cisco. Potrebno je vnesti zemljevide prostorov, nanje postaviti dostopne točke. Omogoča tudi kalibracijo, s katero se izboljša natančnost določanja lokacije. Podjetje Ruckus Wireless je v ta namen razvilo aplikacijo za pametne mobilne telefone, s katero je možno natančno premeriti signale. Zato sklepamo, da za določanje lokacije uporabljajo

¹Ruckus zahteva uporabo posebnih dostopnih točk, ki so nekoliko dražje od navadnih, sicer uporaba storitve *SPoT* ni mogoča.

podoben algoritem, ki se uporablja v metodi RADAR. Do lokacij mobilnih naprav je možno dostopati preko naprednega vmesnika REST, ki omogoča dogodke *push*. SPoT ponuja tudi analizo podatkov, ki pa ni tako zmogljiva kot pri Cisco MSE, ampak je primerljivejša z Meraki Presence Analytics.

4.4 Primerjava

V tabeli 4.1 primerjamo vse tri produkte med seboj glede na ceno, programski vmesnik (API), kalibracijo, analitiko, natančnost, enostavnost in ali se storitev uporablja v oblaku ali ne.

	cena	API	kalibracija	analitika	natančnost	oblak	enostavnost
Cisco	-	+	da	+	+	ne	-
Meraki	+	o	ne	+	-	da	++
Ruckus	o	++	da	o	+	da	o

Tabela 4.1: Primerjava produktov za določanje lokacije z Wi-Fi. V primeru, da smo lastnost označili z *++*, potem ta storitev zares izstopa v primerjavi z ostalima dvema.

Cisco in Ruckus ponujata zelo podobni rešitvi. Obe sta v osnovni namenjeni velikim podjetjem, ki z brezžičnim omrežjem Wi-Fi pokrivajo velike površine. Ti produkti predstavljajo dodano vrednost, saj na eni strani omogočajo navigiranje v notranjih prostorih, po drugi stran pa tudi analitiko, s katero podjetja pridobijo vpogled, kaj se dogaja v bližini dostopnih točk. Če pa želimo določiti lokacije mobilnih naprav z visoko natančnostjo, potem je potrebno opraviti dolgotrajen postopek kalibracije, kar seveda pordraži storitev. Po drugi strani Meraki ponuja dober kompromis, med ceno, enostavnostjo in analitiko, a je kvaliteta določanja lokacije veliko slabša.

Poglavje 5

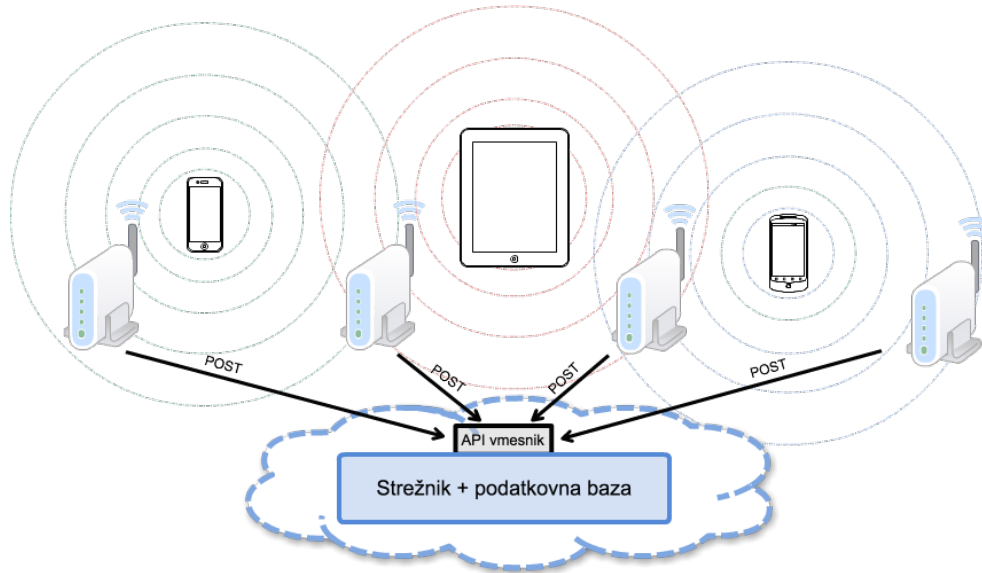
Razvoj sistema za zajem signalov Wi-Fi

V tem poglavju predstavimo vse pomembne gradnike, s katerimi smo zgradili sistem za zajem signalov Wi-Fi. Sprva predstavimo arhitekturo sistema in obe glavni komponenti: dostopne točke ter strežnik za shranjevanje in obdelavo podatkov. Uvedemo dve obliki predstavitve podatkov: intervale prisotnosti in vektorski opis signala. Opišemo težave, na katere smo naleteli pri razvoju in predstavimo rešitve, s katerimi smo zagotovili daljši čas neprekinjenega delovanja. Razložimo, kako se podatki prenašajo med dostopnimi točkami in strežnikom ter kako strežnik pripravi podatke za nadaljnjo analizo. Opišemo tudi zajem vremenskih podatkov in proces zajema prometa Wi-Fi s programom Wireshark oz. tshark. Ob koncu poglavja na kratko opišemo še problem zasebnosti zaradi naslovov MAC in naš predlog.

5.1 Arhitektura sistema

Sistem za zajem signalov Wi-Fi je sestavljen iz dveh glavnih komponent: dostopnih točk, ki skrbijo za zajem signala Wi-Fi in strežnika, ki skrbi za hrambo podatkov, zajem vremenskih podatkov in pripravo podatkov za nadaljnjo analizo. Shema je predstavljena na sliki 5.1. V nadaljevanju podrob-

neje predstavimo sestavo dostopnih točk in vlogo strežnika.



Slika 5.1: Arhitektura sistema za zajem signalov Wi-Fi

5.1.1 Dostopne točke

Ker signal Wi-Fi potuje prosto po prostoru, ga je možno enostavno zaznati z navadnimi brezžičnimi karticami, ki so prisotne v vsakem sodobnem računalniku. Eden izmed ciljev magistrske naloge je vzpostavitev čim cenejšega sistema, zato bomo za zajem uporabili cenovno ugodne mikroračunalnike Raspberry Pi (slika 5.2) in dodatne USB brezžične kartice Wi-Pi (slika 5.3). Trenutna cena računalnika Raspberry Pi znaša okoli 25 EUR, cena brezžične kartice pa 7 EUR. Osnovna cena dostopne točke (Raspberry Pi + 2x brezžična kartica Wi-Pi) tako znaša manj kot 40 EUR, medtem ko lahko za dostopne točke opisane v 4. poglavju odštejemo 500 EUR in več.

Zakaj potrebujemo dve brezžični kartici? Če želimo zajemati ves brezžičen promet, potem je potrebno eno kartico nastaviti v promiskuiteten način (več o tem v razdelku 5.3). Ker pa kartica v tem načinu ne more vzpostaviti povezave z internetom, potrebujemo dodatno brezžično kartico, ki skrbi le za

komunikacijo z internetom. V primeru, da je na mestu priključitve dostopne točke na voljo žična povezava, jo uporabimo za komunikacijo z internetom, saj ima Pi že vgrajen priključek Ethernet. Tako je potrebna le še ena brezžična kartica, kar še dodatno poceni dostopno točko.

Na Raspberry Pi lahko naložimo poljuben operacijski sistem UNIX. Mi smo naložili Raspbian. Ta izhaja iz družine operacijskih sistemov Debian ter velja za preverjen in stabilen sistem.

Ob zagonu operacijski sistem samodejno zažene skripto, ki skrbi za zagon programa Wireshark (natančnejši proces je opisan v razdelku 5.3), obdelavo zajetih podatkov in komunikacijo s strežnikom. Vsi zajeti paketi, ki niso *probe request*, se zavržejo. Za vsak *probe request* pa se zabeležijo naslednji podatki:

- časovni žig (angl. **timestamp**), ki označuje čas oddaje signala;
- naslov MAC naprave, ki je poslala paket;
- moč signala v dBm.



Slika 5.2: Raspberry Pi, model B. Vir slike: *Wikipedia*

Nato skripta kreira zahtevek POST in na strežnik pošlje zgornje tri podatke, poleg pa doda še svoj naslov MAC. Na podlagi slednjega lahko strežnik ugotovi, katera dostopna točka je zaznala paket.

Dostopna točka ima lokalni pomnilnik. V primeru, da izgubi povezljivost s strežnikom, podatke shranjuje v lokalni pomnilnik. Ko se povezava izgubi, se požene diagnostični test in preveri, kaj je šlo narobe. Najprej preveri gonilnike Wi-Pi kartice, če še delajo pravilno in poizkuša odpraviti napako. Nato preveri prostor na SD kartici in, v primeru pomanjkanja praznega prostora, počisti `\tmp` mapo. Če napak ne more odpraviti, po dveh minutah ponovno zažene operacijski sistem. Ko se povezava spet vzpostavi, se prenesejo vsi podatki na strežnik. Na ta način preprečimo izgubo podatkov in povečamo zanesljivost sistema.



Slika 5.3: Wi-Fi brezžična omrežna kartica za Raspberry Pi = Wi-Pi.

Kljub temu, da danes računalniki Raspberry Pi veljajo za stabilne sisteme, smo naleteli na naslednje nevšečnosti:

Slaba združljivost SD kartic: Naleteli smo na številne SD kartice, ki niso združljive s strojno opremo računalnika Raspberry Pi. V primeru uporabe nezdružljive kartice se težave pokažejo že takoj ob zagonu, saj se računalnik ne zažene. Rešitev: nakup SD kartice, ki je uradno podprta in obstaja na seznamu uradno potrjenih kartic [4].

Nezanesljivi napajalniki: V šest-mesečnem obdobju testiranja, so odpovedali trije napajalniki. Rešitev: nakup boljših in močnejših napajalnikov (2500 mA), kar pa dostopno točko podraži za dodatnih 5 EUR. Osnovni napajalniki zmorejo le 1200 mA, kar je premalo za normalno delovanje, saj dodatni dve brezžični kartici zahtevata veliko energije.

Stabilnost Wi-Pi kartice: pogosto prihaja do sesutja gonilnikov, ki poganjajo Wi-Pi brezžično kartico. Rešitev: v primeru, da Wireshark

dlje časa ne zazna Wi-Fi prometa, se zaženejo diagnostični testi, ki poskušajo odpraviti napako. Skrajna rešitev je ponovni zagon operacijskega sistema.

Kljub nevšečnostim so dostopne točke s postopnimi izboljšavami skozi čas postale zelo stabilne. Brez težav delajo neprekinjeno tudi več kot mesec dni, če v tem obdobju seveda ne zmanjka električne energije.

5.1.2 Strežnik

Druga glavna komponenta, ki sestavlja sistem za zajem podatkov Wi-Fi, je strežnik, ki vsebuje podatkovno bazo SQL, REST vmesnik in logiko za pretvarjanje podatkov. Glavne naloge, ki jih opravlja, so naslednje:

Zbiranje podatkov iz dostopnih točk: Strežnik prek vmesnika REST sprejema zahteve POST, ki jih pošiljajo dostopne točke. Zahteve obdela in podatke o probe requestih shrani v bazo. Za vsak *probe request* se zapiše: zgoščen naslov MAC naprave, naslov MAC dostopne točke, moč signala [dBm], časovni žig paketa in trenutna lokacija dostopne točke¹.

Zbiranje vremenskih podatkov: Strežnik skrbi tudi za zajem vremenskih podatkov. Na spletu obstajajo številni programski vmesniki (angl. Application Interface oz. API), ki omogočajo dostop do vremenskih podatkov. Izbrali smo spletno storitev, ki jo ponuja *World Weather Online* [29]. Glavna prednost te storitve je ta, da ponuja tudi dostop do 60-dnevne zgodovine. Strežnik enkrat dnevno prenese vremensko napoved in jo shrani v bazo.

Preoblikovanje podatkov za nadaljnjo analizo: S strežnika pridobimo podatke z zahtevki GET. Vsak zahtevek GET mora vsebovati tudi dva parametra: časovno obdobje in format podatkov (intervali prisotnosti

¹Strežnik mora biti vedno seznanjen z lokacijo, kje se v danem trenutku dostopne točke nahajajo.

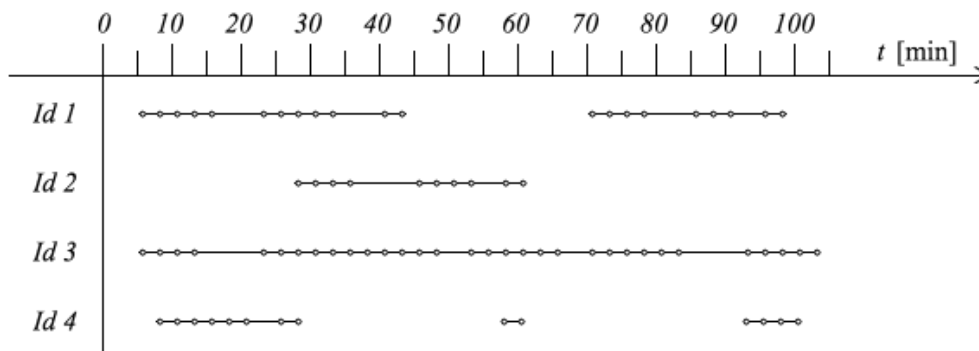
oz. vektorski opis zaznanega signala). Preden strežnik vrne odgovor, se podatki ustrezno prevedejo glede na izbrani format.

5.2 Predstavitev podatkov

Definirali smo dve obliki predstavitve podatkov, s katerima želimo poenostaviti nadaljnjo analizo. Prva oblika, intervali prisotnosti, je namenjena podatkom, ki jih pridobimo, če imamo na voljo le eno dostopno točko. Druga, vektorski opis signala, pa je namenjena predstavitvi podatkov zajetih z več dostopnimi točkami. V nadaljevanju bomo obe obliki podrobno opisali.

5.2.1 Intervali prisotnosti

Za vsako napravo lahko seznam zaznanih probe requestov pretvorimo v obliko intervala prisotnosti. Ta predstavitev podatkov omogoča enostaven izračun grafov (več v razdelku 8.1). Laična definicija: interval prisotnosti predstavlja prisotnost telefona v bližini dostopne točke in je definiran s časovnima žigoma, ki označujeta začetek in konec prisotnosti.



Slika 5.4: Pretvorba seznama probe requestov v intervale prisotnosti za 4 različne naprave. Mali krogec predstavlja probe request, krogi povezani s črto pa intervale prisotnosti.

Definicija intervala prisotnosti: Interval prisotnosti se začne s časovnim žigom prvega probe requesta, pred katerim že več kot X časa dostopna točka

ni zaznala drugega in se zaključi s časovnim žigom zadnjega probe requesta, za katerim ni bil zaznan drug probe request vsaj X časa.

V definiciji se nahaja parameter X , s katerim določimo najmanjšo časovno razliko med dvema časovnima intervaloma. Na sliki 5.4 je prikazana pre-tvorba iz seznama probe requestov v intervale prisotnosti. V tabeli 5.1 so naštetih intervali.

naprava	intervali
Id 1	[5, 44], [71, 98]
Id 2	[28, 61]
Id 3	[5, 103]
Id 4	[7, 29], [58, 61], [94, 103]

Tabela 5.1: Vsak interval je predstavljen z začetnim in končnim časovnim žigom (v tem primeru v minutah).

5.2.2 Vektorski opis signala

Sistem za zajem signala Wi-Fi sestavlja N dostopnih točk. Signal, ki ga odda naprava, lahko zazna več dostopnih točk. Za predstavitev bomo uporabili vektorje, definirane z N komponentami. Vsaka komponenta predstavlja moč signala, ki jo je izmerila dostopna točka za isti signal.

Poglejmo si primer sistema s štirimi dostopnimi točkami (DT_1, DT_2, DT_3, DT_4), torej $N = 4$. Neka naprava je oddala 5 probe requestov in vsak je predstavljen s časovnim žigom $(t_1, t_2, t_3, t_4, t_5)$. Podatki so predstavljeni v tabeli 5.2.

Pomembno je, da ohranjamo vrstni red komponent. Če poznamo moč signala, lahko izračunamo tudi razdaljo (z enačbo 3.10) ali kvaliteto signala (z enačbo 3.11). Ta format zapisa podatkov predstavlja vstopno točko za klasifikacijo s trilateracijo (več v razdelku 6.3) in klasifikacijo naprav med prostori s strojnim učenjem (več v razdelku 6.4).

čas	DT_1 [dBm]	DT_2 [dBm]	DT_3 [dBm]	DT_4 [dBm]
t_1	-54	-72	-	-84
t_2	-	-66	-	-80
t_3	-59	-63	-70	-84
t_4	-51	-73	-61	-82
t_5	-52	-77	-61	-
\vdots	\vdots	\vdots	\vdots	\vdots

Tabela 5.2: Vektorski opis signala za 5 različnih časovnih žigov. Pomišljaj (-) pomeni, da dostopna točka (DT) ni zaznala signala.

5.3 Zajemanje signala Wi-Fi z orodjem Wireshark

Wireshark je odprtokodno orodje, ki omogoča zajemanje mrežnih paketov. Podprt je na vseh treh glavnih platformah (Windows, MacOS, Linux) in se uporablja za odkrivanje težav v omrežjih ter v izobraževalne namene. Razširjeno je tudi orodje *tcpdump*, ki je že privzeto nameščeno v operacijskih sistemih Unix.

Wireshark lahko nastavi mrežno kartico v promiskuiteten način (angl. promiscuous mode), kar omogoča pregled vsega prometa, ki pride do te mrežne kartice. To pomeni, da kartica obdrži vse pakete² in s tem lahko analiziramo tudi promet, ki ni namenjen tej mrežni kartici. Monitor oz. RFMON (Radio Frequency MONitor) način pa omogoča zajemanje prometa tudi takrat, ko kartica ni povezana v omrežje.

Za zajem probe requestov z orodjem Wireshark, je potrebno izvesti naslednje tri korake:

Zagon programa Wireshark: Program je potrebno zagnati z uporabnikom, ki ima dovolj pravic, da lahko spreminja nastavitve na brezžični

²V splošnem kartice delujejo tako, da pakete, ki niso naslovljeni na to mrežno kartico, zavrzijo. Govorimo o t.i. ne-promiskuitetnem načinu (angl. non-promiscuous mode).

kartici.

Vklop promiskuitetnega načina: V nastavitvah programa je potrebno vključiti promiskuiteten način.

Vklop filtra: Ker nas zanimajo le probe requesti, vključimo filter:

$$wlan.fc.type == 0 \ \&\& \ wlan.fc.subtype == 4$$

Prvi del filtra se omeji samo na pakete tipa management frames³, drugi del pa omeji prikaz samo na probe requeste⁴.

Glede na zgoraj opisani postopek, smo tudi sami pognali program Wireshark in preučili vrnjene podatke. Na sliki 5.5 so prikazani zajeti probe requesti. Poglejmo si поближе, katere pomembnejše podatke vsebuje posamezen probe request:

pošiljateljev naslov MAC - Wireshark prvih 24 bitov avtomatično prevede v ime proizvajalca. Uporablja mehanizem, ki smo ga predstavili v razdelku 2.1;

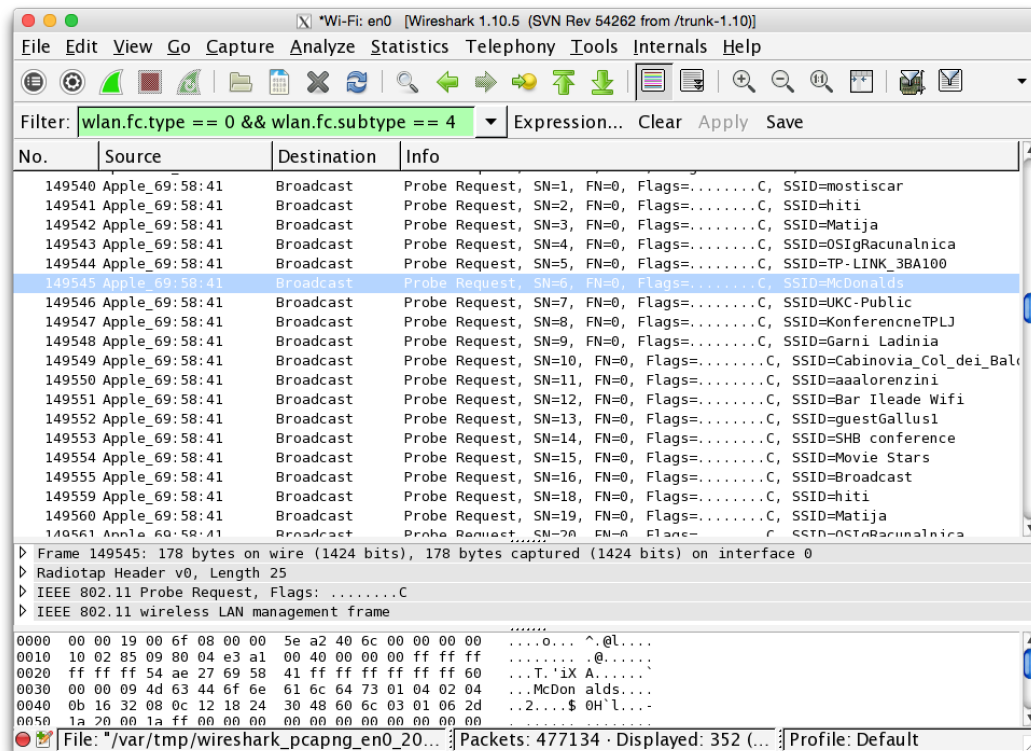
prejemnikov naslov MAC - uporablja se broadcast naslov *ff:ff:ff:ff:ff:ff*;

SSID - V tem polju naprava, ki pošilja probe request, določi ime brezžičnega omrežja, kateremu želi poslati paket. V primeru, da želi probe request poslati vsem napravam, nastavi SSID polje na *broadcast*. Obstajata namreč dve vrsti probe requestov: unicast in broadcast, več v razdelku 2.1;

epoch time/unix time - časovni žig, ki označuje čas odpošiljanja. Predstavlja število milisekund, ki so minile od 1.1.1970;

³Standard 802.11 loči med tremi različnimi tipi okvirjev: management (0), control (1) in data (2) okvirji.

⁴Obstaja več različnih tipov management okvirjev: association request (0), association response (1), probe request (4), probe response (5), beacon (8), authentication (11), deauthentication (12).



Slika 5.5: Zajeti paketi z orodjem Wireshark. Prikazane pakete je poslal iPhone 5S ob ponovnem zagonu vmesnika Wi-Fi.

SSI Signal - moč signala, ki ga je izmerila brezžična kartica ob zajemu signala. Ta podatek se nahaja v *radiotap* glavi, ki ga naknadno doda brezžična kartica. Podrobnejši opis se nahaja v razdelku 3.2. Podatek je predstavljen z enoto dBm.

SSI Noise - nivo šuma ob anteni brezžične kartice. Podobno kot SSI Signal, tudi ta podatek doda brezžična kartica ob zajemu signala. Več v razdelku 3.2.

Poleg orodja Wireshark, ki ima grafični vmesnik, obstaja tudi orodje tshark [38]. Ta se izvaja v ukazni vrstici, kar omogoča enostavno uporabo v programski kodi. Z naslednjim ukazom poženemo zajemanje prometa na vmesniku *wlan0*:


```
tshark -i wlan0 -I -f 'broadcast' -R 'wlan.fc.type == 0 && wlan.fc.subtype  
== 4' -T fields -e frame.time_epoch -e wlan.sa -e radiotap.dbm_antsignal
```

Promet s filtrirom (-f) omejimo le na probe requeste in izpisujemo naslednja tri polja: časovni žig paketa (*time_epoch*), pošiljatelj naslov MAC (*wlan.sa*) in moč signala (*dbm_antsignal*). Ta ukaz predstavlja ključni del sistema za zajem podatkov Wi-Fi. Uporabljajo ga dostopne točke. Podrobnejši opis se nahaja v razdelku 5.1.1.

5.4 Problem zasebnosti

V razdelku 2.1 smo predstavili postopek vzpostavljanja povezave med napravo in dostopno točko. Obstajata dva načina: pasivni in aktivni, predvsem slednji pa je problematičen s stališča zasebnosti uporabnika mobilne naprave. Pri aktivnem iskanju se namreč pošiljajo paketi, ki oglašujejo naslov MAC od naprave in pa imena brezžičnih omrežij (SSID), na katera se je naprava v preteklosti uspešno povezala. SSID-ji pogosto vsebujejo pomenljiva imena, npr. ime bara, restavracije, trgovine, šole, ipd., na podlagi katerih lahko sklepamo, kje se je mobilna naprava v preteklosti nahajala. V raziskavi, ki je bila objavljena v članku [11] pa so naredili še korak več. Na podlagi SSID-jev, so raziskovalci uspeli zelo natančno ugotoviti, iz katere države uporabnik prihaja in kateri jezik govori. V okviru te magistrske naloge, SSID-jev nismo potrebovali, zato smo jih ob zajemu vsakega paketa, zavrgli.

Drugače pa je z naslovi MAC, saj zaradi narave problema, ki ga rešujemo, naslov MAC nujno potrebujemo. Ta služi kot identifikator, na podlagi katerega lahko ločimo obiskovalce med seboj. V zadnjih letih države spreminjajo zakone na področju informacijske zasebnosti, ki postajajo vse strožji. Predvsem v Evropski uniji imajo posamezne države zelo stroge zakonodaje, med katere spada tudi Slovenija. Trenutno slovenska zakonodaja obravnava naslov MAC kot osebni podatek, kar pomeni, da je potrebno biti pri rokovanju s tovrstnimi podatki še posebej previden. Ena izmed rešitev, ki jo

uporabljajo tudi produkti, predstavljeni v 4. poglavju, je zgoščevanje naslovov MAC. Na spletu smo odkrili več člankov, ki so pri zgoščevanju naslovov MAC odkrili ranljivosti. Doktorski študent Univerze Stanford, Jonathan Mayer, je s pomočjo oblačne platforme AWS, uspel izračunati naslove MAC iz zgoščenih vrednosti zgoščevalne funkcije SHA1 le v 12 minutah [18]. Zato smo v okviru magistrske naloge vse naslove MAC pretvorili v zgoščene vrednosti z zgoščevalno funkcijo SHA256, ki je zaradi svoje dolžine varnejša v primeru napada z grobo silo. Poleg naslova MAC, smo za vhodni podatek uporabili še datum in ime lokacije. S tem smo onemogočili povezovanje podatkov o istih napravah preko več dni in več lokacij.

$$SHA256(datum||lokacija||naslov\ MAC) \quad (5.1)$$

Problematike oglaševanja naslova MAC se vse bolj zavedajo tudi proizvajalci mobilnih naprav. Apple je prvi, ki poizkuša ta problem rešiti in sicer z generiranjem naključnih naslovov MAC v času, ko mobilna naprava išče brezžična omrežja. Prvo rešitev so predstavili ob izidu nove verzije operacijskega sistema iOS 8, septembra 2014. Kmalu po izidu so omrežni strokovnjaki preverili, kako se ta rešitev obnese v praksi [12]. Ugotovili so, da iPhone generira naključne naslove MAC samo takrat, ko so izpolnjeni naslednji pogoji:

- telefon mora imeti vključen način spanja, oz. zaslon mora biti izključen (angl. sleep mode);
- vmesnik Wi-Fi mora biti vključen, vendar povezava z brezžičnim omrežjem ne sme biti vzpostavljena;
- storitev *Location services* mora biti izključena;
- podatkovni prenos mora biti izključen;
- telefon mora biti iPhone 5S oz. novejši.

Če so ti pogoji izpolnjeni, potem iPhone zares generira naključne naslove MAC. V tem primeru, se naslovi spremenijo vsakič, ko uporabnik prižge in

ugasne zaslon. Zaenkrat le malo telefonov generira naključne naslove MAC, saj so pogoji prestrogi. V prihodnosti pa lahko pričakujemo, da bo algoritem izpopolnjen in da bodo podobno rešitev vpeljali tudi ostali proizvajalci. Takrat bo sledenje mobilnim napravam oteženo, še vedno pa bo mogoče relativno štetje mobilnih naprav v bližini dostopnih točk.

Pri implementaciji sistema za zajem signalov Wi-Fi smo spoštovali zasebnost uporabnikov. SSID-jev nismo shranjevali, vse naslove MAC pa smo ustrezno zgostili po formuli 5.1.

Poglavje 6

Klasifikacija signalov Wi-Fi v regije

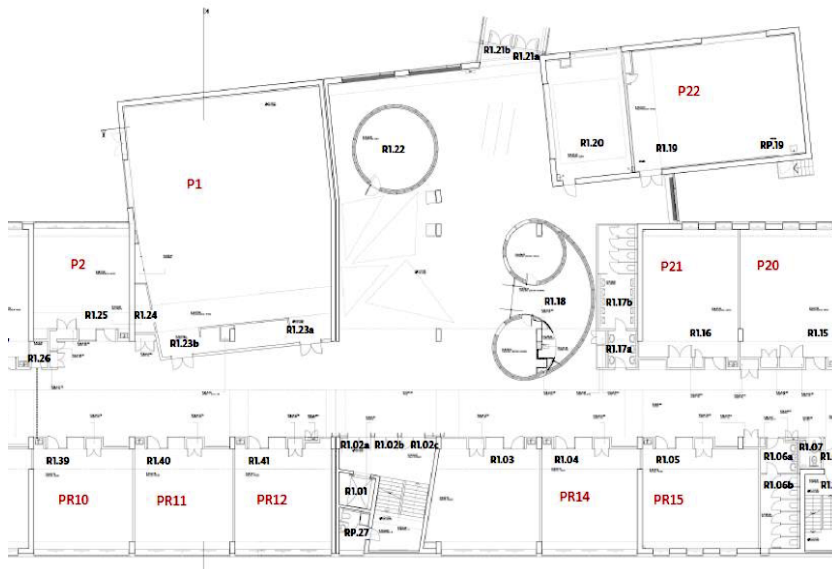
Cilj magistrske naloge je vzpostaviti enostaven in poceni sistem za zajem signalov Wi-Fi. V skladu s tema dvema ciljema smo zgradili sistem, ki smo ga podrobneje opisali v 5. poglavju. Pri dostopnih točkah smo uporabili poceni antene Wi-Fi, ki se po kvaliteti ne morejo primerjati z antenami, uporabljenimi v današnjih komercialnih produktih. Podrobneje smo jih predstavili v 4. poglavju. Ti produkti uporabljajo antene z natančnejšimi merilnimi instrumenti za določanje moči signala, kar omogoča boljše lociranje. V razdelku 3.2 smo pisali o odbojih elektromagnetnega valovanja v notranjih prostorih in kako težko je pretvoriti moč signala v razdaljo. Zato dražje dostopne točke z boljšimi antenami in boljšimi merilnimi instrumenti pripomorejo k natančnejšemu lociranju, kar pa je pravzaprav tudi njihova primarna naloga.

V magistrski nalogi pa primarna naloga sistema Wi-Fi ni določanje natančne lokacije, ampak štetje obiskovalcev po prostorih. Zato lahko problem določanja lokacije poenostavimo in poskušamo določiti le regije oz. prostore, v katerih so se mobilne naprave nahajale.

6.1 Definicija problema

Problem klasifikacije signalov Wi-Fi v regije lahko prevedemo v enostaven matematičen problem z množicami. Vhodni podatki so naslednji:

Tloris zgradbe: Primer tlorisa zgradbe je predstavljen na sliki 6.1. Potreben je predvsem za boljšo predstavo o tem, kako so razporejeni prostori, hkrati pa služi kot osnova za definicijo regij. Ta vhodni podatek ni nujno potreben.



Slika 6.1: Primer tlorisa zgradbe.

Seznam regij: S seznamom regij določimo interesne regije/prostore, ki nas zanimajo pri analizi. Dovolj je, da regije pisno opišemo (npr. Predavalnica 1, Stopnišče, Hodnik 1, WC) in si zapomnimo, kako so regije razporejene. Alternativno lahko uporabimo tloris zgradbe in nanj vrišemo poligone, ki predstavljajo regije. Zaželeno je, da meje regij potekajo po stenah prostorov. S tem izkoristimo fizikalne zakonitosti, saj vsaka stena močno vpliva na elektromagnetno valovanje. Pri seznamu

regij je potrebno upoštevati le en pogoj: regije se med seboj ne smejo prekrivati.

Mat. definicija: Naj bo R množica regij r_i , ker $i = 1, \dots, n$ (n predstavlja število regij). Pogoj, da se regije med seboj ne smejo prekrivati, je izpolnjen natanko takrat, ko velja:

$$\bigcap_{i=1}^n r_i = \emptyset, r_i \in R \quad (6.1)$$

Vektorski opis signalov: V razdelku 5.2.2 smo uvedli t.i. vektorski opis signalov. Vektor predstavlja opis zaznanega signala Wi-Fi. V tabeli 6.1 vsaka komponenta predstavlja zaznano moč signala ob določeni dostopni točki. Te vrednosti lahko poljubno pretvorimo v razdalje z enačbo 3.10 oz. v kakovost signala s pomočjo enačbe 3.11.

Mat. definicija: Naj bo \vec{s} vektor opisan z N komponentami:

$$\vec{s} = (s_1, s_2, s_3, \dots, s_N), \quad (6.2)$$

kjer i -ta komponenta predstavlja moč signala izmerjenega ob dostopni točki DT_i in je N število dostopnih točk. Potem je S množica vseh signalov zapisanih v vektorski obliki.

ID	DT_1 [dBm]	DT_2 [dBm]	DT_3 [dBm]	DT_4 [dBm]	regija
id_1	-54	-72	-	-84	?
id_2	-	-66	-	-80	?
id_3	-59	-63	-70	-84	?
id_4	-51	-73	-61	-82	?
id_5	-52	-77	-61	-	?
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

Tabela 6.1: Primer vektorskega opisa signalov z neznanimi regijami.

Definirajmo klasifikacijo kot funkcijo K , ki vektorju \vec{s} priredi regijo $r_i \in R$. V primeru, da nobena regija ne ustreza tem pogojem, potem naj funkcija vrne *neznana regija*, kar pomeni, da je signal izviral v nedefinirani regiji.

$$K(\vec{s}) = r_i \in R \vee \text{neznana regija} \quad (6.3)$$

V nadaljevanju predstavimo algoritme, s katerimi lahko vsakemu vektorju \vec{s} določimo pripadajočo regijo $r_i \in R$ in tako ugotovimo, v katerem prostoru je bila naprava, ko je oddala signal.

V 3. poglavju smo predstavili tri algoritme, s katerimi lahko izračunamo lokacijo signala.

Metoda najmočnejše dostopne točke vrne za dani vektor \vec{s} lokacijo najmočnejše dostopne točke. Posledično vrača le koordinate dostopnih točk. To pomeni, da lahko kreiramo največ N regij (toliko kot je dostopnih točk) in vsaka regija mora vsebovati vsaj eno dostopno točko. Metodo podrobneje predstavimo v razdelku 6.2.

Profiliranje signala Wi-Fi (angl. finger-printing) je najboljša metoda za določanje lokacije, katere pomanjkljivost je veliko število začetnih meritev oz. kalibracija, ki zahteva veliko časa. Z uporabo te metode ne izpolnimo cilja magistrske naloge, ki zahteva čim hitrejšo in enostavno vzpostavitev sistema.

Trilateracija je glede na cilje magistrske naloge najprimernejša metoda, saj je enostavna za vzpostavitev. Podrobneje jo predstavimo v razdelku 6.3.

6.2 Klasifikacija z metodo najmočnejše dostopne točke

Metoda najmočnejše dostopne točke za dani vektor \vec{s} , vrne lokacijo najmočnejše dostopne točke. Če želimo signale klasificirati po regijah, potrebujemo presli-

kavo, ki lokacijo posamezne dostopne točke DT_i preslika v regijo r_j . Klasifikacija je možna le, če imamo na voljo več dostopnih točk od regij: $|S_{DT}| \geq |R|$. Idealno pa je, če je ta preslikava bijektivna. V tem primeru se mora v vsaki regiji nahajati natanko ena dostopna točka. Metoda dela dobro tudi v primeru, če imamo na voljo več dostopnih točk kot regij, vendar pa v tem primeru sistem še dodatno podražimo. Sledi primer preslikave, ki vsako dostopno točko preslika v eno izmed regij:

najmočnejša DT	regija
$DT_1 \implies$	regija 1
$DT_2 \implies$	regija 2
$DT_3 \implies$	regija 3
$DT_4 \implies$	regija 4
\vdots	\vdots

Pri definiciji preslikave se bomo držali načela, da lokacijo dostopne točke preslika v tisto regijo, v kateri se nahaja. V primeru, da dostopna točka leži izven vseh definiranih regij, potem naj slika v najbližjo regijo, v kateri ni druge dostopne točke.

6.3 Klasifikacija s trilateracijo

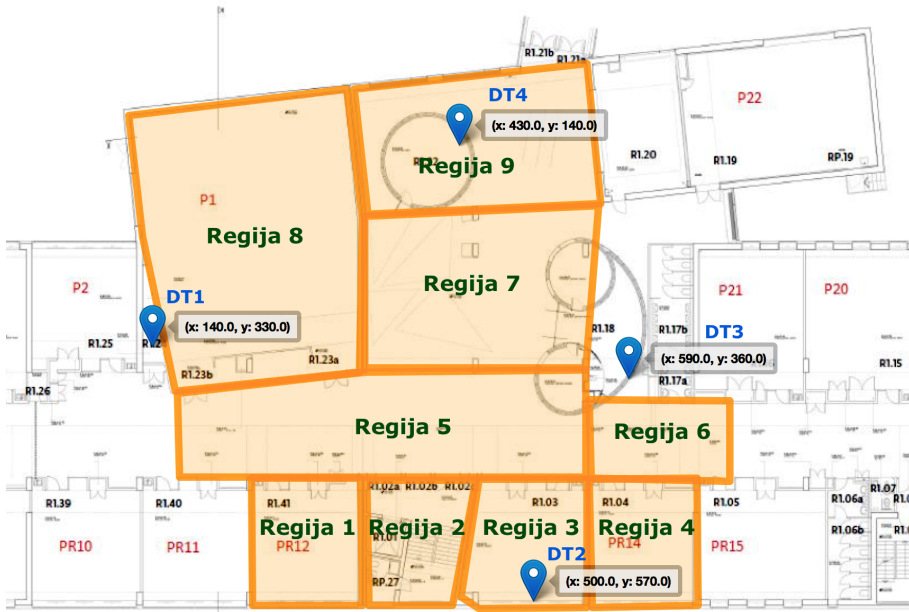
Za izvajanje algoritma trilateracije je potrebno:

Vpeljati koordinatni sistem: To lahko naredimo tako, da vzamemo sliko tlorisa in definiramo zgornji levi kot za izhodišče koordinatnega sistema (0,0), spodnji desni kot pa naj predstavlja koordinata (širina slike, višina slike). Pri tem lahko uporabimo poljubno mersko enoto, npr. slikovne pike ali pa metre.

Označiti lokacije dostopnih točk: Za vsako dostopno točko DT_i je potrebno določiti lokacijo (x_i, y_i) na tlorisu. Seznam dostopnih točk definirajmo z množico: $S_{DT} = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$.

Opisati regije s poligoni: Vsako regijo je potrebno opisati kot poligon:

$r_i = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots\}$, ki na tlorisu predstavlja regijo v realnem prostoru. Pomembno je, da se poligoni med seboj ne prekrivajo.



Slika 6.2: Tloris zgradbe z regijami in lokacijami štirih dostopnih točk.

Šele nato se lahko lotimo klasifikacijo z metodo trilateracije, ki smo jo opisali v razdelku 3.6. Definirali smo jo kot funkcijo $tril()$, ki vektorju \vec{s} in seznamu dostopnih točk S_{DT} priredi koordinato (x, y) :

$$tril(\vec{s}, S_{DT}) = (x, y) \quad (6.4)$$

Za regijo $r_i \in R$, v kateri je naprava oddala signal \vec{s} , mora veljati naslednji pogoj:

$$r_i \Leftrightarrow tril(\vec{s}, S_{DT}) = (x, y) \subset r_i, r_i \in R \quad (6.5)$$

V primeru, da trilateracija vrne točko, ki ne leži v nobenem izmed poligonov, ki določajo regije, potem se signal klasificira kot *neznana regija*.

Na primeru si pogledimo določitev regije vektorju \vec{s} . Vzemimo tloris zgradbe, ki je predstavljen na sliki 6.1. Najprej definirajmo koordinatni sistem. Levi zgornji kot slike naj predstavlja izhodišče $(0, 0)$. Širina slike meri $862px$, višina pa $594px$. Spodnji desni kot naj torej predstavlja koordinato $(862, 594)$. Določimo koordinate še za dostopne točke. Označili smo lokacije štirih dostopnih točk: $DT_1(140, 330)$, $DT_2(430, 140)$, $DT_3(590, 360)$, $DT_4(500, 570)$. Na koncu definiramo še poligone za vsako regije. Dan je primer poligona za regijo 8: $\{(114, 109), (326, 81), (341, 203), (339, 350), (161, 373), (127, 217)\}$

Rezultat je prikazan na sliki 6.2. Sedaj lahko za vsak vektor določimo pripadajočo regijo. Tabela 6.2 prikazuje izračunane regije vektorjev po zgoraj opisanem postopku.

ID	DT_1 [dBm]	DT_2 [dBm]	DT_3 [dBm]	DT_4 [dBm]	regija
id_1	-54	-72	-	-84	Regija 7
id_2	-	-66	-	-80	Regija 7
id_3	-51	-90	-50	-51	Regija 3
id_4	-51	-	-61	-90	Regija 9
id_5	-52	-77	-61	-	Regija 3

Tabela 6.2: Rezultat klasifikacije s trilateracijo.

6.4 Klasifikacija s strojnim učenjem

Strojno učenje (angl. machine learning) je eno izmed področij računalništva, ki s pomočjo posebnih algoritmov išče zakonitosti v podatkih. Deli se na več različnih področij, eno izmed njih pa je tudi klasifikacija. To področje se ukvarja s klasifikacijo podatkov v ločene razrede. Sprva s pomočjo testnih podatkov, ki vsebujejo tudi uvrstitev v razrede, poišče povezave med podatki in razredi, te pa nato združi v model, na podlagi katerega lahko razvrsti tudi podatke z neznanim razredom.

V primeru klasifikacije signalov Wi-Fi v regije, lahko za vhodne podatke

vzamemo zaznane signale, za razrede pa regije. Vsak zaznan signal je uvrščen v tisto regijo, v kateri je signal izviral. Zaznane signale bomo opisali s pomočjo vektorskega opisa signala, ki smo ga definirali v razdelku 5.2.2. Primer testne množice podatkov, je prikazan v tabeli 6.1.

DT_1 [dBm]	DT_2 [dBm]	DT_3 [dBm]	DT_4 [dBm]	regija
-54	-72	-	-84	regija 1
-51	-66	-	-80	regija 1
-	-73	-61	-82	regija 2
-52	-	-61	-	regija 3
-52	-77	-61	-	regija 4
\vdots	\vdots	\vdots	\vdots	\vdots

Tabela 6.3: Vektorski opis signalov in znane regije.

Da lahko klasificiramo signale v regije, potrebujemo testno množico podatkov. Opraviti moramo torej meritve signalov za vsako regijo. To lahko naredimo tako, da se za nekaj časa postavimo v vsako regijo, zabeležimo čas pričetka/konca merjenja, med tem časom pa hodimo naokoli po regiji in z mobilno napravo oddajamo signale Wi-Fi. Z enostavnim trikom lahko ta postopek pospešimo. Če na mobilni napravi hitro prižigamo in ugašamo vmesnik Wi-Fi, naprava hitreje oddaja pakete *probe request*. Sistem za zajem signalov Wi-Fi bo zaznal signale in ker vemo, kdaj smo bili v kakšnem prostoru, lahko enostavno povežemo signale z regijami. Dobro je, da opravimo čim več meritev in da pokrijemo tudi skrite kotičke regij. Le tako lahko zgradimo kvaliteten model, ki bo omogočal dobro klasifikacijo.

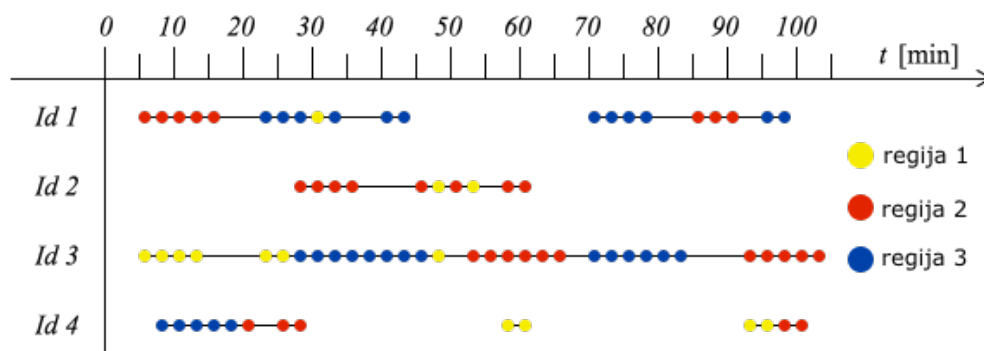
S pomočjo teh meritev, pridobimo testno množico podatkov, s katero lahko z različnimi klasifikatorji zgradimo model, ki omogoča klasifikacijo med regijami. Obstaja veliko različnih klasifikatorjev, npr. SVM, naivni Bayes, logistična regresija, K najbližjih sosedov, itd. Zaradi obsežnosti magistrske naloge, ne bomo šli v podrobnosti, kako posamezen klasifikator deluje. Dober opis oz. razlaga vseh, pa se nahaja v knjigi *Data Science for Business* [30].

Zaenkrat smo izmerjene signale predstavili z močjo signala, zaznanega ob posamezni dostopni točki. V 3. poglavju smo podrobneje pregledali, kako lahko moč signala pretvorimo v razdaljo oz. kvaliteto signala. Za razdaljo lahko uporabimo formulo 3.10, za kvaliteto signala pa enačbo 3.11.

V razdelku 9.4 preverimo na realnih podatkih, kateri vhodni podatki in kateri klasifikatorji se obnesejo najboljše.

6.5 Od klasifikacije do obiska po regijah

V razdelku 5.2.1 smo uvedli intervale prisotnosti, ki smo jih zgradili na podlagi zaznanih signalov Wi-Fi, konkretnije paketov *probe request*. Podobno lahko naredimo z vektorji signalov. Za vsak vektor lahko s pomočjo klasifikacije napovemo regijo, v kateri se je nahajala naprava v času oddaje signala.



Slika 6.3: Pretvorba seznama vektorjev v intervale prisotnosti za 4 različne naprave. Mali krogec predstavlja vektor, krogi povezani s črto pa intervale prisotnosti. Barva predstavlja regijo, v katero ga je razvrstila klasifikacija.

Na sliki 6.3 je prikazan primer za 4 naprave. Vsak krogec predstavlja vektor, ki opisuje zaznano moč signala za vsako dostopno točko. Barva kroga predstavlja regijo, v katero je posamezen vektor uvrstila klasifikacija.

Na ta način lahko vidimo, kako se je naprava premikala skozi čas. Ker pa klasifikacija ni 100 % pravilna, se večkrat zgodi, da naprava na videz “skače” med regijami. To težavo pa lahko odpravimo z medianinim filtrom.

Poglavje 7

Postavitev sistema in statistika zajetih podatkov

V 5. poglavju smo opisali razvoj sistema za zajem signala Wi-Fi. Sistem smo sprva postavili v stanovanjsko hišo, kjer smo testirali stabilnost delovanja. Po mesecih razvoja in testiranja je sistem postal dovolj stabilen za zajem velikih količin podatkov. Kljub majhnemu številu obiskovalcev stanovanjske hiše, smo uspeli zaznati nočno/dnevne vzorce. Zaznali smo tudi številne šibke signale, ki so jih najverjetneje oddajali mobilni telefoni pešcev. Pojavi se potreba po klasifikaciji signalov, zato smo pričeli z iskanjem najboljšega načina za razvrščanje signalov med prostori. Algoritme smo natančneje opisali v 6. poglavju. Prvi rezultati so bili zelo spodbudni. Da pa bi lahko potrdili oz. znanstveno izmerili, kako dobro deluje sistem in algoritmi za razvrščanje, smo potrebovali večji nabor podatkov.

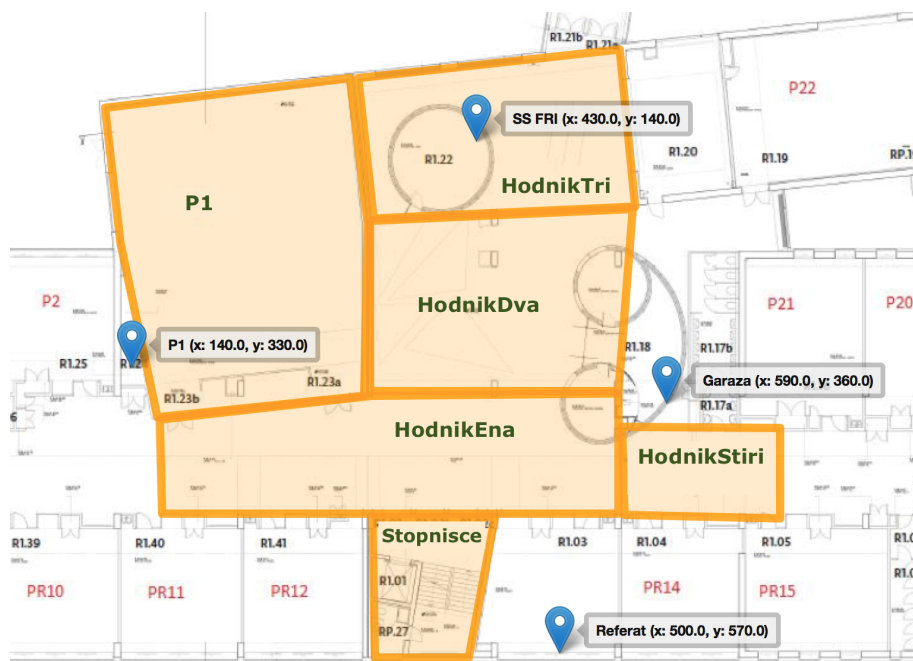
Sistem smo zato prenesli na Fakulteto za računalništvo in informatiko, v Ljubljani (v nadaljevanju FRI). Fakulteto vsakodnevno obišče veliko študentov, profesorjev in ostalih obiskovalcev. Ker gre za tehnično podkovane ljudi, sklepamo, da obiskovalci FRI pogosteje uporabljajo pametne mobilne telefone od ostalih ljudi. Upravičeno lahko predpostavljamo, da bo podatkov glede na število obiskovalcev več, kot pa če bi sistem postavili na kakšno drugo javno ustanovo. Kljub temu sklepamo, da vzorec ljudi ni povsem irelevanten, saj

glede na vzroke, ki smo jih našli v uvodu 2. poglavja, lahko v prihodnje pričakujemo, da bo vse več ljudi uporabljalo pametne prenosne naprave.

Urnik predavanj na FRI je javno dostopen na spletni strani [42], poleg tega so objavljeno tudi vsi pomembnejši dogodki, ki se odvijajo znotraj fakultete. Ti podatki nam omogočajo enostavno iskanje korelacije med dogodki in zajetimi podatki.

V tem poglavju bomo predstavili shemo postavitve in statistično analizo podatkov, ki smo jih zajeli v dobrih dveh mesecih. Podrobnejša analiza teh podatkov se nadaljuje v 8. poglavju.

7.1 Shema postavitve



Slika 7.1: Razporeditev regij in dostopnih točk na Fakulteti za računalništvo in informatiko

Podatke smo zajemali v pritličju FRI, natančneje med predavalnico P1, referatom, garažo in prostori Študentskega sveta. Razporeditev dostopnih

točk in regij je predstavljena na sliki 7.1. Dostopne točke smo postavili na štiri lokacije:

P1 - Dostopno točko smo postavili v majhen kabinet poleg predavalnice P1, ki ga obdajajo betonske stene, s predavalnico pa ga povezujejo za signale Wi-Fi prepustna vrata.

Študentski svet FRI - Dostopno točko smo pritrdili pod mizo pisarne Študentskega sveta. Stene okoli pisarne so lesene in dobro prepuščajo signale Wi-Fi.

Garaža - Dostopno točko smo postavili na tla ob stekleno steno, ki delno ovira signal Wi-Fi. Prostor obdajajo tudi lesene stene, ki dobro prepuščajo signal.

Referat - Dostopna točka je bila postavljena na okenski polici. Prostor referata so obdani z betonskimi stenami, ki močno omejujejo signal Wi-Fi.

Določili smo tudi šest približno enako velikih regij. S tem smo izničili vpliv velikosti, ki lahko vpliva na rezultate klasifikacije. Definirali smo dva različna tipa regij:

Tip 1: regija se ujema s prostorom, definiranim z betonskimi stenami;

Tip 2: regija predstavlja del večjega prostora in se delno ujema s stenami.

P1 in *Stopnisce* sta regiji tipa 1, regije *HodnikEna*, *HodnikDva*, *HodnikTri*, *HodnikStiri* pa tipa 2. Slednje predstavljajo velik osrednji prostor fakultete. Z razbitjem na manjše enote smo uspeli ohraniti enake velikosti regij. Hkrati pa smo z izbiro regij, ki med seboj niso ločene s stenami, omogočili eksperiment, s katerim bomo preverili, kako dobro dela klasifikacija, če uporabljamo le regije tipa 1 in kako, če imamo mešane tipe.

Eden izmed ciljev magistrske naloge je tudi čim nižja cena sistema. Pri tem smo poskušali uporabiti cenovno ugodnejše komponente, ki pa še vedno dovolj dobro opravljajo svojo nalogo. Izračunajmo ceno testnega sistema:

komponenta	cena	količina	skupaj
Raspberry Pi	25 EUR	4	100 EUR
Wi-Fi	7 EUR	6	42 EUR
SD kartica	10 EUR	4	40 EUR
Polnilec	6 EUR	4	24 EUR
			206 EUR

Cena uporabljenega sistema je bila 206 EUR. Ocenjujemo, da bodo ti stroški že v bližnji prihodnosti še krepko nižji. Pojavljajo se novi projekti, npr. *C.H.I.P.: The World's First Nine Dollar Computer* [14], ki prinaša majhne računalnike z integriranim vmesnikom Wi-Fi za le 9 dolarjev (okoli 8 EUR). Ko bo tak računalnik na voljo, bo potrebno za vzpostavitev enakovrednega sistema plačati le 20 % današnje cene.

7.2 Statistika zajetih podatkov

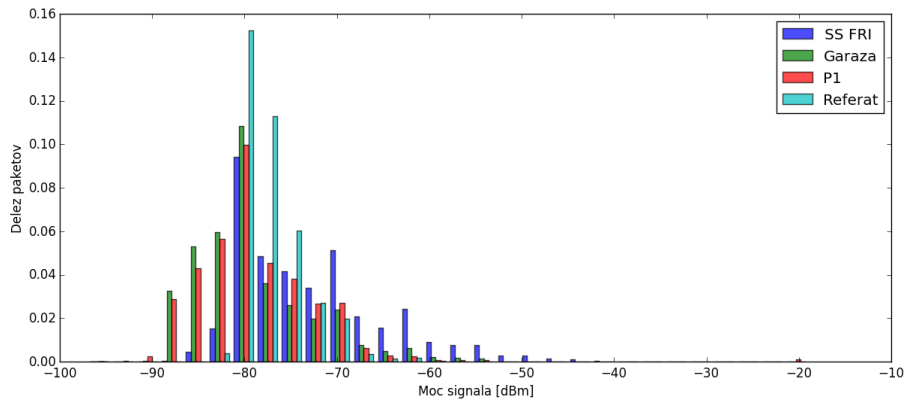
Podatke smo zbirali v obdobju med 13. marcem in 12. majem 2015. V tem času smo s štirimi dostopnimi točkami zbrali 6.362.843 paketov *probe request*. V tabeli 7.1 je podrobneje predstavljena statistika za vsako dostopno točko (DT) posebej. Vidimo lahko, da sta dostopni točki DT_{SSFRI} in DT_{Garaza} zaznali precej več prometa v primerjavi z DT_{P1} in $DT_{Referat}$. Predvsem slednja je zaznala trikrat manj prometa kot DT_{Garaza} . Rezultat ni presenetljiv, saj sta bili prvi dve dostopni točki postavljeni v hodniku, kjer je bilo veliko obiskovalcev, hkrati pa sta bili obdani tudi z lesenimi stenami, ki manj ovirajo elektromagnetno valovanje v primerjavi z betonskimi stenami.

Za lažjo predstavbo smo izračunali povprečno število zajetih paketov na časovne enote. V povprečju so dostopne točke vsako minuto zaznale več kot 70 probe requestov. Ponoči in ob vikendih prometa skoraj ni bilo in če vzamemo v obzir le delovne dneve v času med 8.00 in 17.00, smo v povprečju zaznali več kot 300 probe requestov na minuto. Pošiljali smo jih na virtualni strežnik z 1 GB delovnega pomnilnika in 20 GB SSD prostora, ki s procesiranjem podatkov ni imel težav. Ker kompleksnost procesiranja z dodajanjem

	DT_{SSFRI}	DT_{Garaza}	DT_{P1}	$DT_{Referat}$	Skupaj
vseh skupaj	1.972.024	2.142.308	1.573.521	596.193	6.362.843
na dan	32.867,1	35.705,1	26.225,3	9.936,5	106.047,4
na uro	1.369,5	1.487,7	1.092,7	414,0	4.418,6
na minuto	22,8	24,8	18,2	6,9	73,6

Tabela 7.1: Število zajetih probe requestov po dostopnih točkah glede na časovno obdobje.

dostopnih točk teče linearno $O(N)$, pri čemer N predstavlja število dostopnih točk, in se podatki lahko zapisujejo vzporedno, lahko sistem enostavno horizontalno skaliramo z dodajanjem strežnikov. Predvidevamo, da je arhitektura sistema, ki smo jo predstavili v razdelku 5.1, primerna tudi za velike sisteme s 100 in več dostopnimi točkami.



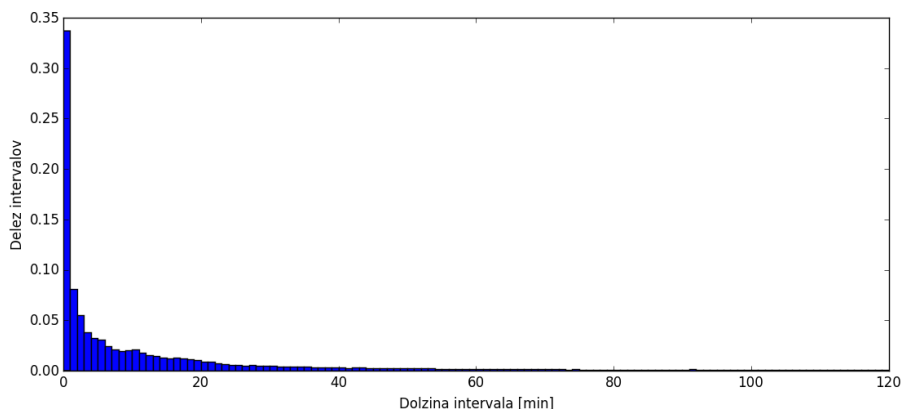
Slika 7.2: Primerjava porazdelitve moči signalov med vsemi štirimi dostopnimi točkami.

Na sliki 7.2 so predstavljene porazdelitve moči signalov, ki so jih zaznale posamezne točke. Vse štiri dostopne točke so zaznale signale z močjo med -50 in -90 dBm, daleč največji delež pa predstavljajo signali moči -80 dBm. Dostopna točka DT_{SSFRI} je v povprečju zajela najmočnejše signale, verjetno zato, ker je bila postavljena v prostor z lesenimi stenami, ki zelo dobro prepuščajo elektromagnetno valovanje. Odgovorov, zakaj je porazdelitev signalov takšna kot je, še nimamo. Sklepamo pa, da na rezultate močno vpliva

razporeditev prostorov.

7.2.1 Statistika intervalov prisotnosti

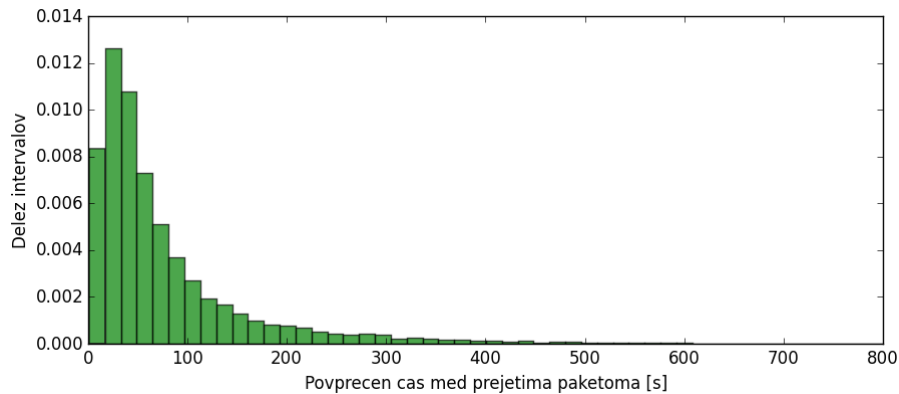
Intervale prisotnosti smo definirali v razdelku 5.2.1. V obdobju testiranja smo našli 148.539 intervalov prisotnosti dolgih vsaj 1 sekundo. Narisali smo histogram 7.4, ki prikazuje porazdelitev intervalov glede na dolžino. Vidimo lahko, da je veliko intervalov zelo kratkih (mimobežniki), nekaj je tudi zelo dolgih (statične naprave). Večina intervalov pa je krajših od 20 min.



Slika 7.3: Histogram dolžin intervalov prisotnosti.

S pomočjo intervalov prisotnosti smo izračunali tudi podatek o tem, kako pogosto telefoni oddajajo probe requeste. V literaturi smo sicer večkrat zasledili, da jih mobilne naprave oddajajo vsakih 15-60 sekund [11], vendar smo to želeli tudi sami preveriti. Kako smo izmerili ta podatek? Vzeli smo vse intervale prisotnosti daljše od 1 minute. Nato smo v vsakem intervalu izračunali povprečen čas med dvema probe requestoma, pri tem pa smo zanemarili tiste, ki so prišli skoraj ob istem času (± 1 milisekunda). Nato smo narisali histogram 7.4, ki prikazuje porazdelitev intervalov glede na povprečen čas med dvema probe requestoma.

Vidimo lahko, da znaša čas v povprečju med dvema probe requestoma pri večini intervalov manj kot minuto. Mediana znaša 48 sekund, aritmetična



Slika 7.4: Histogram intervalov prisotnosti glede na povprečen čas med dvema probe requestoma.

sredina pa 76 sekund. Rezultat je torej zelo podoben navedenemu v strokovnem članku [11].

7.2.2 Statistika vektorjev

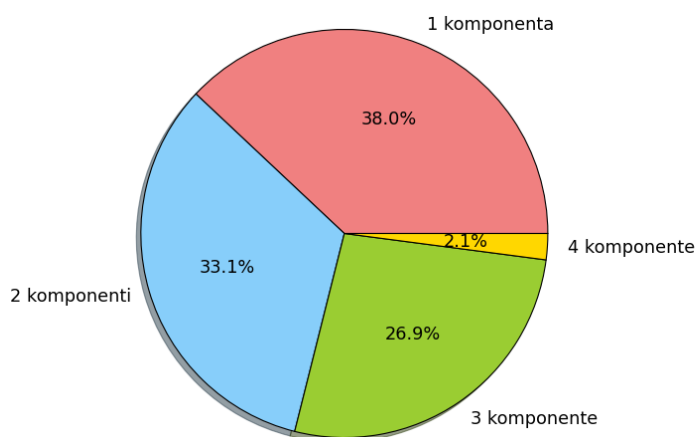
S pomočjo vektorjev smo opisali signale, ki so jih zaznale štiri dostopne točke: DT_{SSFRI} , DT_{Garaza} , DT_{P1} in $DT_{Referat}$. Vsaka komponenta vektorja predstavlja moč signala zaznanega ob posamezni dostopni točki. Vektorje smo razdelili v dve skupini:

vektorji z znano regijo - vemo, v kateri regiji se je nahajala naprava, ko je oddala signal;

vektorji z neznano regijo - vsi preostali vektorji, za katere ne poznamo regije.

Vektorje z znano regijo smo pridobili z merjenjem signalov. Meritve smo opravljali dvakrat, 26.3. in 13.4.2015 in sicer zato, da bi preverili, če se signal skozi čas spreminja. Bistvenih razlik med dnevoma nismo zaznali. Podatke Wi-Fi smo generirali z mobilnim telefonom iPhone 5S. Postavili smo se v vsako regijo in s pomočjo spletnega vmesnika (predstavljen je v dodatku A.6) zapisali začetni in končen čas merjenja za izbrano regijo. Med tem

časom smo generirali signale Wi-Fi tako, da smo večkrat v minuti vključili in izključili vmesnik Wi-Fi. S tem smo spodbudili telefon, da je oddal čim več signalov v kratkem času. Skupaj smo v dobrih 10 minutah pridobili 242 različnih vektorjev z znanimi regijami, podrobneje pa so predstavljeni v dodatku C. Sledi kratka statistična analiza zajetih vektorjev.

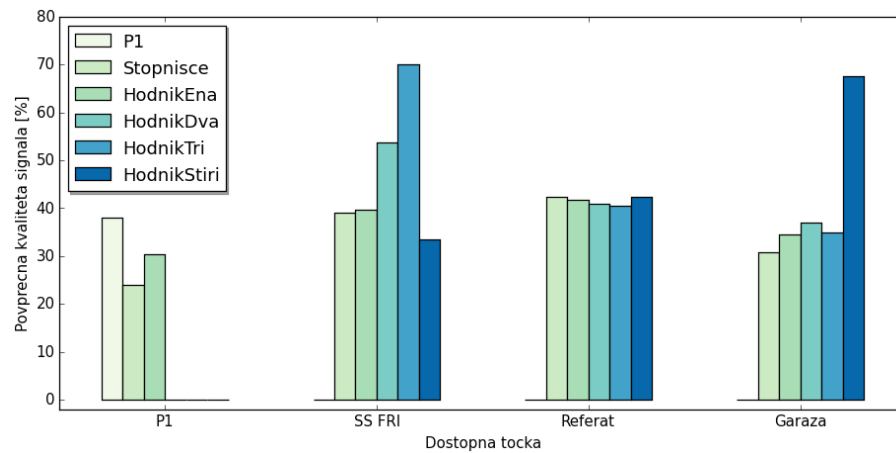


Slika 7.5: Deleži glede na število definiranih komponent vektorjev.

Zanimalo nas je, koliko dostopnih točk je hkrati zaznalo posamezen signal Wi-Fi oz., koliko vektorjev ima definirano 1, 2, 3 oz. 4 komponente. Rezultati so vidni na sliki 7.5. Več kot $\frac{3}{4}$ vektorjev ima definirano le eno oz. dve komponenti. Kaj to pomeni? Če ima vektor le eno komponento, potem metoda trilateracije vrača enake rezultate kot metoda najmočnejše točke. V primeru dveh komponent, lahko s trilateracijo določimo lokacijo le na daljici, ki povezuje obe dostopni točki. Približno četrtnina vektorjev pa ima 3 oz. 4 komponente. Ti vektorji pa predstavljajo kvalitetne podatke, katerih bi želeli imeti čim več.

Zanimalo nas je še, kakšne signale so zaznale dostopne točke glede na regijo. Najprej smo vse moči signalov pretvorili v kvaliteto signala 3.11, saj si to mero lažje predstavljamo. Nato smo izračunali povprečen signal po regijah za vsako dostopno točko in narisali graf, ki je predstavljen na sliki 7.6.

Analiza po dostopnih točkah:



Slika 7.6: Povprečna kvaliteta signala ob dostopni točki glede na regije. Prazen stolpec pomeni, da dostopna točka ni zaznala signalov za to regijo.

DT_{P1} je zaznala signale le s treh najbližjih regij: *P1*, *Stopnisce* in *HodnikEna* in še ti so bili slabše kvalitete, glede na ostale dostopne točke. V povprečju so bili signali kvalitete manj kot 40 %, verjetno predvsem na račun dejstva, da je bila dostopna točka v kabinetu poleg predavalnice *P1*, ki ga obdajajo betonske stene.

DT_{SSFRI} je najmočnejše signale zaznala iz regije *HodnikTri*, kar je pričakovano, saj se je dostopna točka nahajala v tej regiji. Močnejši signali so prišli tudi iz regije *HodnikDva*, ki je druga najbližja, vmes pa ni betonskih ovir.

$DT_{Referat}$ je zaznala signale z vseh regij, razen *P1*. V povprečju je zaznala enako kvalitetne signale, ne glede na regijo. To je slabo za klasifikacijo, saj nam ta komponenta nič ne pove. Očitno je bila dostopna točka preveč oddaljena od definiranih regij.

DT_{Garaza} je najmočnejše signale zaznala iz regije *HodnikStiri*, ki se nahaja poleg dostopne točke. Povprečni signali iz ostalih regij, so bili približno enaki.

Torej, že na oko lahko razberemo določene zakonitosti, kar je dober znak za klasifikacijske algoritme. Rezultati pa so tudi pokazali, da dostopna točka postavljena v referatu ne prinaša veliko dodane vrednosti in bi jo bilo treba prestaviti bližje ostalim regijam.

Poglavje 8

Analiza rezultatov

Že v prejšnjem poglavju smo pokazali, da lahko z enostavno statistično analizo razberemo določene zakonitosti. V tem poglavju pa bomo pregledali rezultate, ki jih lahko izračunamo s pomočjo intervalov prisotnosti (več v 5.2.1) oz. vektorskimi opisi signalov (več v 5.2.2). Analizirali bomo podatke predstavljene v prejšnjem poglavju.

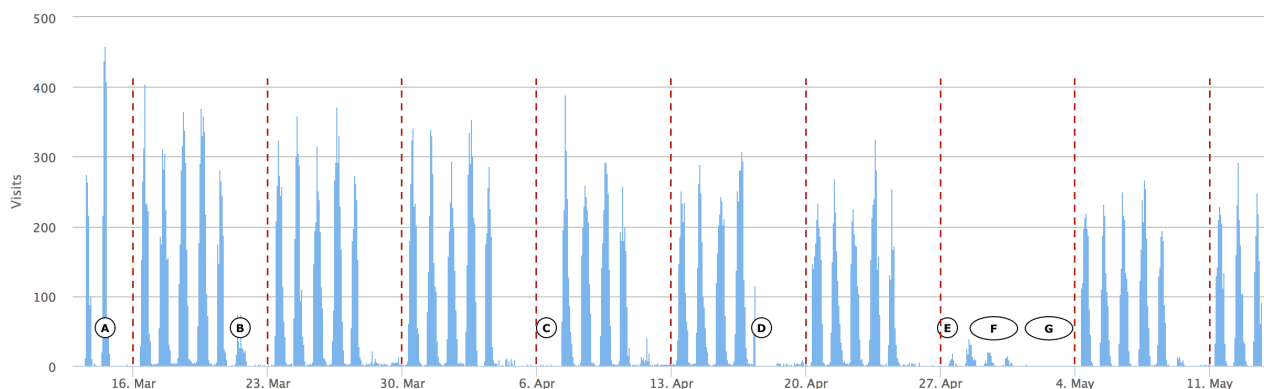
To poglavje je razdeljeno na dva dela: v prvem predstavimo rezultate, ki jih lahko zajamemo z eno dostopno točko (obisk skozi čas, obiskanost po dnevih v tednu, časi zadrževanja, korelacija obiska in vremenu), v drugem delu pa predstavimo praktičen primer uporabe klasifikacije signalov, ki jih zajamemo z več dostopnimi točkami.

8.1 Posamezna dostopna točka

Glavni cilj tega razdelka je potrditi hipotezo, ki pravi, da s pomočjo sledenja Wi-Fi lahko zaznamo število prisotnih ljudi in to že z eno dostopno točko. Pri tem si bomo pomagali z urnikom FRI in javnimi dogodki, ki so bili objavljeni na spletni strani Fakultete. Kljub temu, da smo podatke zajemali s štirimi dostopnimi točkami, bomo v tem razdelku predstavili le rezultate za dostopno točko DT_{SSFRI} , saj so rezultati za ostale dostopne točke podobni.

8.1.1 Obisk skozi čas

S pomočjo intervalov prisotnosti smo izračunali graf obiska skozi čas. Graf sestavljajo stolpci, ki predstavljajo urne intervale. Višina stolpca predstavlja število intervalov prisotnosti, ki se vsaj delno prekrivajo z urnim intervalom tega stolpca. Rezultat je prikazan na sliki 8.1.



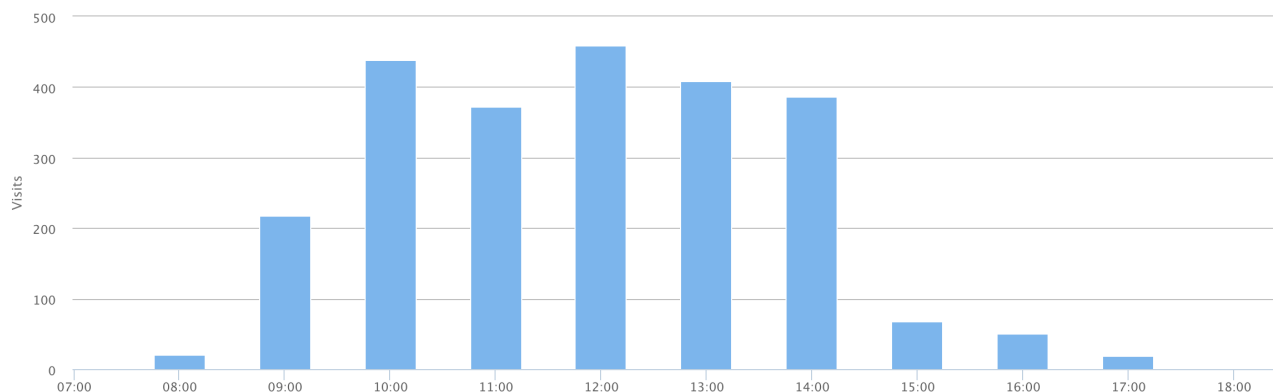
Slika 8.1: Število zaznanih obiskovalcev med 13. marcem in 12. majem. S črkami od A do G so označeni izredni dogodki. Navpične rdeče črtkane črte predstavljajo konec oz. začetek tedna.

Rezultati so na videz zelo prepričljivi. Iz grafa se lepo razberejo delovni dnevi, ko je sistem zaznal več obiskov, in vikendi, ko obiskovalcev skoraj ni moč zaznati. Z rdečo navpično črto smo označili prehode med različnimi tedni¹, črke od A do G pa predstavljajo odstopanja od ustaljenega trenda delovni teden/vikend. Zanimalo nas je, ali lahko razložimo ta odstopanja in s tem potrdimo, da lahko s tovrstnim pristopom res zaznamo relativno število obiskovalcev skozi čas. Sledi pregled vzrokov za ta odstopanja:

A - 14. marca je potekala Web Camp konferenca [3], na kateri je bilo prisotnih okoli 400 tehničnih strokovnjakov. Na sliki 8.2 je prikazan obisk po urah za ta dan. Vidimo lahko, da je bila večina obiskovalcev prisotnih med 10.00 in 15.00. Med tem časom smo v povprečju našli

¹Rdeča navpična črta predstavlja polnoč ob nedeljah.

400 obiskov, kar se zelo lepo ujema z dejanskim številom obiskovalcev. Na to je zagotovo vplivala struktura obiskovalcev, saj gre pretežno za ljudi, ki pogosto uporabljajo Wi-Fi povezavo na svojih pametnih mobilnih telefonih.



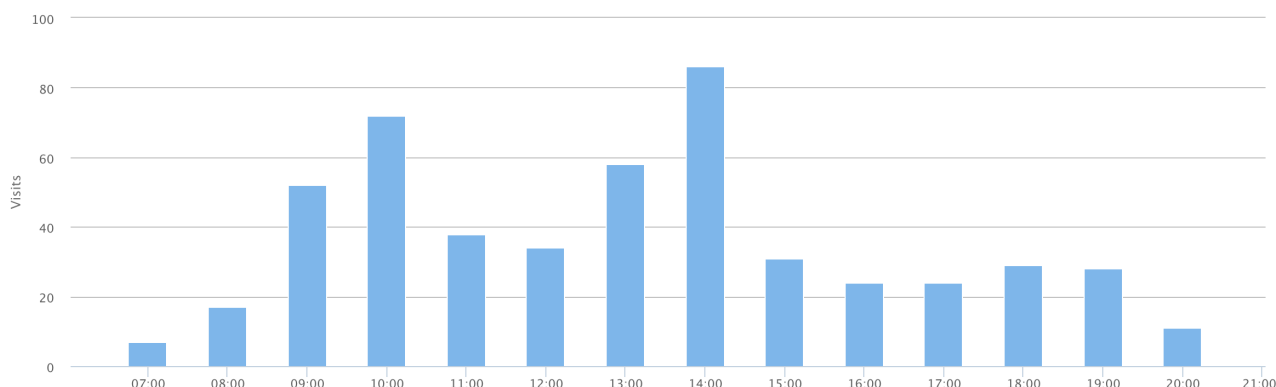
Slika 8.2: Število zaznanih obiskovalcev 14. marca 2015.

Preverimo še ujemanje obiska glede na časovnico dogodka. Registracija prvih udeležencev se je pričela ob 8.00. Ob 9.30 so se pričele prve predstavitve. Med 11.00 in 11.30 je potekal odmor za kavo. Ob 15.00 se je dogodek zaključil. Vidimo lahko, da se časovnica dogodka zelo lepo ujema z grafom na sliki 8.2.

Že drugi dan meritev je pokazal zelo spodbudne rezultate, saj so se izmerjeni podatki zelo dobro ujemali z dejanskim stanjem.

B - 21. marca se je odvijalo tekmovanje v znanju računalništva RTK. Novico o dogodku smo prebrali na spletni strani Fakultete za računalništvo in informatiko [41]. Tekmovanja se je udeležilo 159 srednješolcev iz cele Slovenije [6].

Tekmovanje se je pričelo ob 9.30. Potekalo je v treh skupinah, prvi dve skupini sta imeli na voljo tri ure časa (9.30-12.30), tretja skupina pa pet (9.30-14.30). Razglasitev rezultatov je potekala ob 18.00 [6].



Slika 8.3: Število zaznanih obiskovalcev 21. marca 2015.

Graf na sliki 8.3 prikazuje zaznan obisk po urah. Največ obiskovalcev je sistem zaznal med 9.00 in 11.00 ter 13.00 in 15.00. Ker smo signale Wi-Fi zajemali na hodniku in ne v računalniških učilnicah, kjer je potekalo tekmovanje, se rezultati zdijo pričakovani. Sistem je največ obiskovalcev zaznal pred pričetkom in po koncu tekmovanja. Nekoliko večji obisk smo zabeležili tudi med 18.00 in 20.00, ko je potekala razglasitev rezultatov. Znano je bilo tudi število obiskovalcev (vsaj 159 tekmovalcev [6]). Sistem je hkrati zaznal največ 90 obiskovalcev. Sklepamo, da niso vsi tekmovalci hkrati zaključili z reševanjem nalog. Verjetno tudi ni bila tako visoka penetracija mobilnih telefonov z vključenim signalom Wi-Fi, saj med dijaki omrežje Eduroam ni tako zelo popularno kot med študenti. Kljub vsemu lahko rečemo, da izmerjen obisk sovпада s časovnico dogodka, vendar ne tako zelo prepričljivo, kot npr. v času konference Web Camp.

C - 6. aprila je bil na Velikonočni ponedeljek dela prost dan. Sistem je na uro zaznal največ tri obiske. Sklepamo, da je bila Fakulteta v tem času zaprta. Zaznane signale so lahko poslale obstoječe naprave (pozabljeni mobilni telefoni, tablice, ipd.), ki so tudi sicer ves čas prisotne znotraj Fakultete. Ali pa morda mobilne naprave pešcev, ki so se sprehajali v bližini. Te naprave bi sicer lahko zanemarili, vendar smo želeli na

rezultate čim manj vplivati in s tem pridobiti čim bolj surove podatke. Implementirali smo sicer tudi metodo, ki stacionarne naprave² doda na črno listo (angl. blacklist).

- D - 17. aprila** ob 8.00 je prišlo do popolne odpovedi sistema. Napako smo odpravili 18. aprila ob 10.00. Za to obdobje so se žal vsi podatki izgubili. Odpovedal je glavni strežnik, ki skrbi za shranjevanje podatkov. Strežnik smo gostili na oblaki platformi *DigitalOcean*, ki nudi poceni gostovanje na SSD diskih. Ta dan so izvajali nadgradnjo infrastrukture in med nadgradnjo je prišlo do strežniške napake, ki je povzročila sesutje operacijskega sistema. To je na današnjih oblaknih platformah precej pogost pojav. SLA-ji ne zagotavljajo 100 % dosegljivosti. Da v bodoče ne bi več prišlo do podobnega izpada, smo postavili dodaten strežnik in s tem zagotovili visoko dosegljivost (angl. high availability) našega sistema. Po tem neprijetnem dogodku ni več prišlo do izpada.
- E - 27. aprila** je bila Fakulteta zaradi praznika, Dan boja proti okupatorju, ponovno zaprta. Na ta dan smo zabeležili podobno obnašanje kot na Velikonočni ponedeljek.
- F - 28.,29. in 30. aprila** smo zaznali občutno manj naprav od povprečnega delovnega tedna. Ti trije dnevi so se zvrstili med dvema podaljšanima vikendoma in Univerza v Ljubljani je te tri dni razglasila za študija proste. Kljub vsemu je bila Fakulteta očitno odprta, saj smo zaznali nekje do 15 % povprečnega števila tedenskih obiskovalcev.
- G - 1.,2. in 3. maja** so potekali prvomajski prazniki in FRI je zaprla svoja vrata. Sistem je med tem časom zaznal do 3 naprave na uro, torej v skladu z ostalimi prazniki.

Vidimo lahko, da z enostavno dostopno točko, ki zajema signale Wi-Fi, lahko zelo dobro določimo relativni obisk skozi čas, katerega lahko razložimo

²Def. stacionarna naprave: naprava, ki jo sistem v intervalu dvanaestih ur vidi skupaj več kot deset ur.

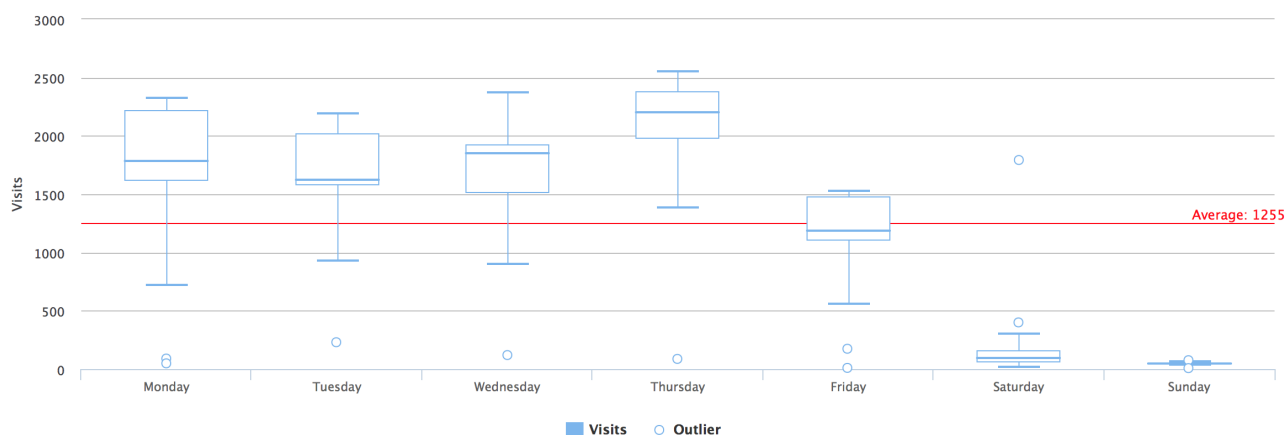
z objektivnimi podatki o tem, kaj se je na Fakulteti v času meritev dogajalo. S tem enostavnim testom smo potrdili, da je s pomočjo sledenja Wi-Fi možno zaznati prisotnost ljudi. Je pa res, da FRI ponuja tudi brezplačno omrežje Eduroam, ki ga uporablja večina obiskovalcev, zato lahko sklepamo, da so zajeti vzorci bližje obnašanju ljudi, kot pa če bi podobne meritve opravljali na kakšni drugi ustanovi.

8.1.2 Obiskanost po dnevih v tednu

V prejšnjem razdelku smo pregledali obiskanost skozi čas. V tem razdelku pa bomo preverili, kako variira obisk glede na dan v tednu. Z grafa na sliki 8.1 lahko čez palec ocenimo, da je obiskovalcev čez vikend veliko manj kot pa med tednom. Razberemo lahko tudi to, da so ponedeljki bolj obiskani od petka, kar smo nekako tudi pričakovali. V nadaljevanju bomo s statistično analizo poskušali dokazati, da so petki manj obiskani, hkrati pa bomo preverili, kateri dan v tednu velja za najbolj obiskanega.

Za statistično analizo po dnevih smo uporabili škatlasti diagram (angl. box and whisker diagram), kot ga je definiral ameriški matematik John Tukey leta 1977 [31]. Za vsak dan med 13.3. in 12.5.2015 smo imeli na voljo 8 meritev, saj je med tem časom minilo 8 tednov. Za vsak dan posebej smo izračunali mediano (srednja črta znotraj škatle), spodnji kvartil Q_1 (spodnja črta škatle), zgornji kvartil Q_3 (zgornja črta škatle) in vrednost $IQR = Q_3 - Q_1$, ki predstavlja ravno velikost škatle. Zgornjo mejo smo določili s formulo $\min(\max Obisk, mediana + 1,5 * IQR)$, spodnjo pa $\max(\min Obisk, mediana - 1,5 * IQR)$. Vrednosti, ki padejo izven spodnje oz. zgornje meje imenujemo osamelci (angl. outliers). S pomočjo teh izračunov smo izrisali diagram na sliki 8.4.

Statistična analiza obiska po dnevih s pomočjo škatlastega diagrama je vrnila pričakovane rezultate. Vikendi so bili manj obiskani od delovnih dni, predvsem v nedeljo tako rekoč nismo zabeležili obiskovalcev. Največja disperzija obiska je prisotna ob ponedeljkih, od torka do petka pa je manjša. Najmanj obiskovalcev med tednom smo po pričakovanjih zabeležili v petek,



Slika 8.4: Obisk po dnevih za obdobje med 13.3. in 12.5.2015.

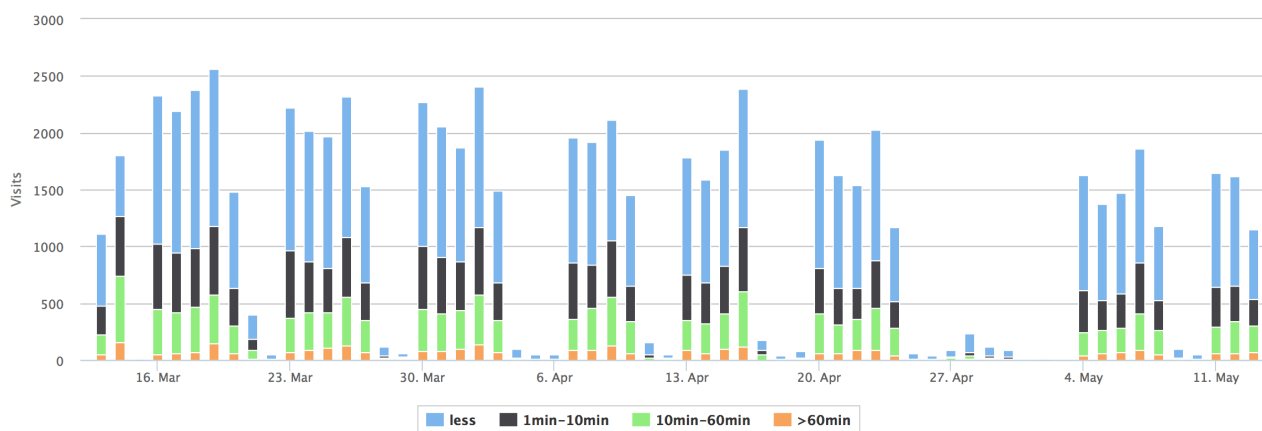
kar ni presenetljivo, saj to razkrije tudi pogled na urnik (v petek je manj predavanj kot ostale dni). Pričakovali smo, da bo obiskanost od ponedeljka do četrтка podobna, to so pokazali tudi rezultati. Presenetil pa nas je četrtek, ko smo v povprečju zaznali največ obiskovalcev, skoraj dvakrat več od petka.

Vidimo lahko tudi kar nekaj osamelcev. Ti predstavljajo nenavadne dogodke, ki smo jih opisali v razdelku 8.1.1. V ponedeljek sta prisotna dva osamelca: Velikonočni ponedeljek in Dan boja proti okupatorju. Tri osamelce smo zaznali med torkom in četrtkom - počitnice pred prvomajskimi prazniki. V petek smo zaznali dva osamelca: izpad sistema in Prvi maj. V soboto sta prav tako prisotna dva osamelca: Web Camp konferenca in zaključek tekmovanja RTK. V nedeljo pa smo zaznali tako malo naprav, da tudi tista dva osamelca, ki sta po definiciji prisotna, premalo odstopata od dejanskih meritev in ju lahko zanemarimo.

8.1.3 Časi zadrževanja

S pomočjo intervalov prisotnosti lahko merimo poleg trenutnega števila mobilnih naprav v bližini, tudi dolžino obiska (angl. dwell time analysis). Dostopna točka, ki je zajemala signale se je nahajala v pisarni ŠS FRI, blizu glavnega vhoda in hodnika, v katerem se nahaja največ ljudi v času odmo-

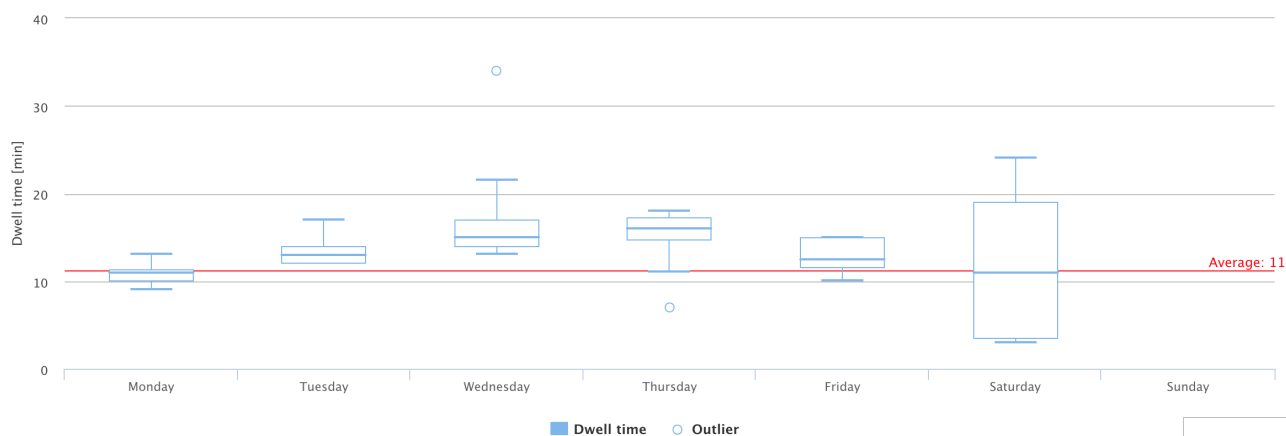
rov. Zato smo pričakovali, da bomo zaznali zelo veliko kratkih intervalov (mimoidoči), veliko srednje dolgih intervalov (obiskovalci med odmori) in nekaj dolgih intervalov (obiskovalci, ki se učijo pri avtomatih). Da bi lahko to hipotezo potrdili, smo obiske razdelili v štiri skupine glede na dolžine intervalov prisotnosti: manj kot minuto, 1-10 minut, 10-60 minut in na več kot eno uro.



Slika 8.5: Časi zadrževanja po dnevih.

Na sliki 8.5 so prikazani rezultati klasifikacije glede na dolžino obiska po dnevih in zelo lepo potrjujejo našo hipotezo. Dobro polovica vseh intervalov je bilo krajših od minute, do 5 % pa daljših od ene ure.

Opravili smo še analizo časov zadrževanja glede na dan v tednu. Rezultat je prikazan na sliki 8.6. Zanemarili smo vse intervale dolžine 0 sekund. Čez teden so povprečni časi zadrževanja podobni, mediane znašajo med 11 in 16 minut. Sklepamo, da je ta čas posledica dolžine odmorov med predavanji, ki znašajo 15 minut. Čez vikend pa so izmerjeni povprečni časi zadrževanja zelo razpršeni, verjetno predvsem na račun manjšega vzorca zajetih podatkov in vpliva statičnih naprav.



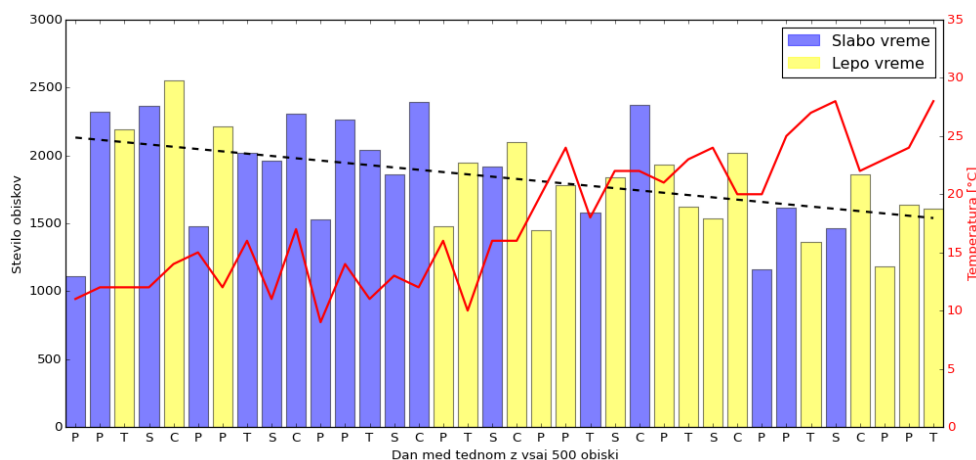
Slika 8.6: Analiza povprečnih časov zadrževanja po dnevih v tednu.

8.1.4 Korelacija obiska in vremena

Poleg zajemanja podatkov Wi-Fi, smo za vsak dan zabeležili še naslednje vremenske podatke³: kratek besedni opis vremena (sončno, oblačno, deževno, ipd.), temperaturo in pritisk. Zanimalo nas je, ali obstaja korelacija med vremenom in obiskom. Ker lahko vreme s pomočjo vremenskih modelov zelo dobro napovedujemo v naprej, bi v primeru korelacije z obiskom lahko napovedali tudi obisk za prihodnost.

Za obdobje med 13. marcem in 12. majem smo za vse dneve med tednom, za katere je sistem zabeležil vsaj 500 obiskov, narisali graf, ki ga prikazuje slika 8.7. Izpustili smo vikende, praznike in prvomajske praznike. Rumena barva predstavlja sončno vreme, modra pa oblačno oz. deževno vreme. Rdeča črta prikazuje trend temperature, črtkana črta pa trend obiska. Vidimo lahko, da s časom temperature rastejo, vreme postaja vse lepše, obisk pa pada. Kljub navidez povezanim rezultatom, pa čiste korelacije med vremenom in obiskom nismo odkrili: npr. v prvem delu smo za tri sončne dneve, resda ob nižji temperaturi, zabeležili nadpovprečen obisk. Z grafa pa lahko razberemo zanimiv trend: ko po daljšem obdobju slabega vremena na vrsto pride lepše vreme, signifikantno upade obisk.

³Za vsak dan smo upoštevali vremenske podatke za 13. uro.



Slika 8.7: Višina stolpca označuje število obiskov, barva predstavlja lepo (sončno) oz. slabo (deževno, oblačno) vreme, rdeča črta predstavlja temperaturo, črtkana siva črta pa trend obiska. Upoštevali smo samo dneve med tednom, ko smo zaznali vsaj 500 obiskov. Črke P, T, S, C, P predstavljajo začetnico imena za dan v tednu.

Zagotovo lahko rečemo, da vreme pomembno vpliva na obisk FRI, vendar pa ni edini dejavnik. Že v razdelku 8.1.2 smo videli, da na obisk močno vpliva dan v tednu. Ker iskanje korelacij podatkov ni primaren cilj te magistrske naloge, podrobne zveze nismo raziskali. Smo pa v dodatek B priložili vse podatke, s pomočjo katerih lahko bralec tudi sam poišče zvezo med obiskom in vremenom.

8.2 Več dostopnih točk

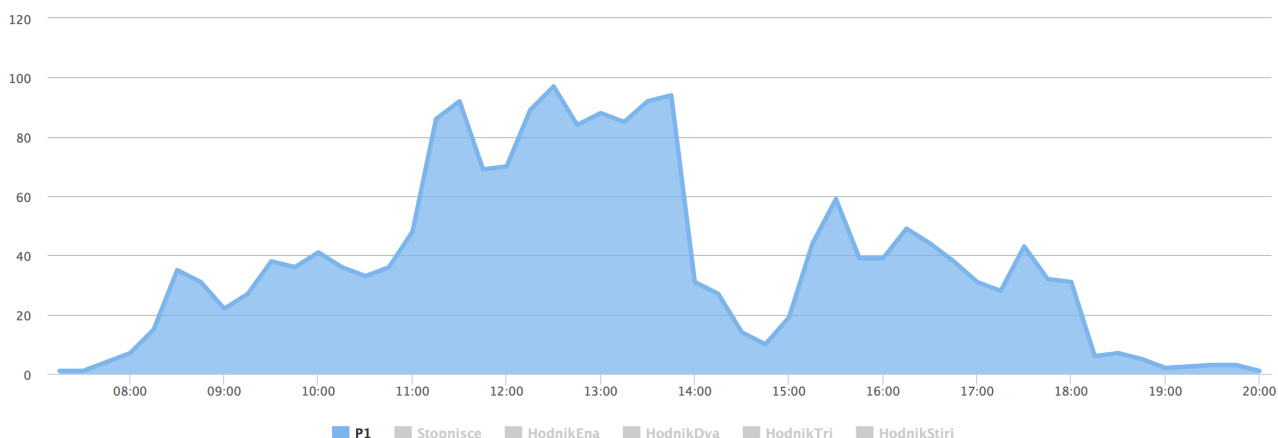
V prejšnjem razdelku smo predstavili rezultate, ki smo jih pridobili s pomočjo le ene dostopne točke. Rezultati potrjujejo, da je možno zaznati relativen obisk blizu dostopne točke skozi čas. Še vedno pa ne vemo, kako blizu so se mobilne naprave nahajale, ko so oddale signal. Lahko da so bile zelo blizu, lahko pa tudi več kot 100 m daleč, v primeru, če je na voljo tako dolga vidna razdalja med napravo in dostopno točko.

Ker pa imamo na voljo več dostopnih točk, lahko oddane signale izmerimo na več znanih lokacijah in na različne načine določimo približno lokacijo izvora signala. V razdelku 6.4 predlagamo klasifikacijo s strojnim učenjem, ki signale Wi-Fi razporedi po prostorih oz. regijah. V nadaljevanju bomo preverili, kako dobro se ta način obnese v praksi.

8.2.1 Zaznavanje predavanj in odmorov

V tem razdelku bomo ovrednotili rezultate klasifikacije na praktičnem primeru. Preverili bomo, kako dobro znamo ločiti med obiskovalci predavalnice P1 in hodnika.

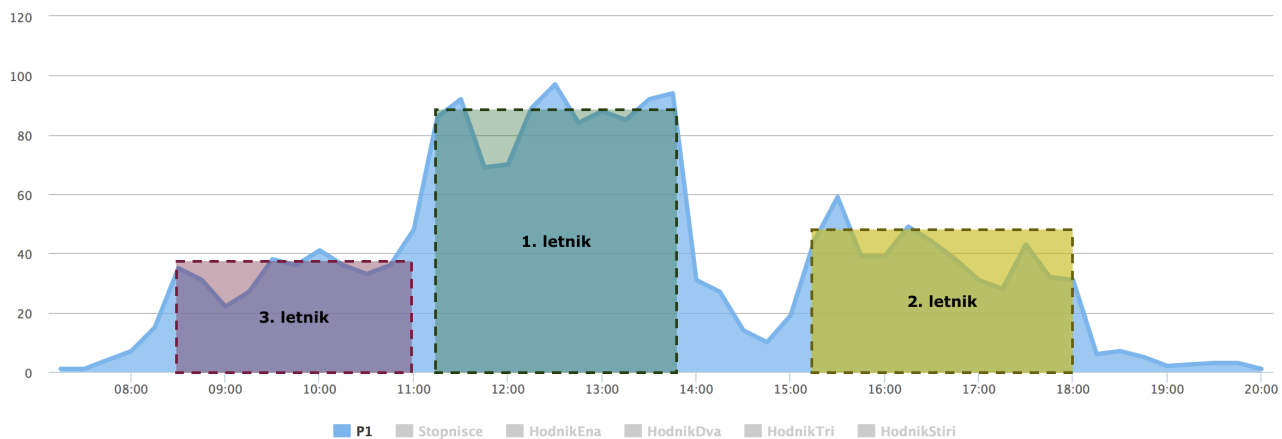
Predstavili bomo meritve obiska po regijah za četrtek, 2.4.2015. Ob četrtkih so na urniku v predavalnici P1 tri predavanja: od 8.00 do 11.00 so predavanja namenjena tretjim letnikom univerzitetnega študija, med 11.00 do 14.00 prvim letnikom in med 15.00 do 18.00 predavanje poslušajo študenti drugega letnika. Glede na urnik, pričakujemo največji obisk v času predavanj, najmanjši pa v času med 14.00 in 15.00, saj je takrat predavalnica, glede na urnik, prazna. Po vsaki polni uri pričakujemo premik študentov iz predavalnice na hodnik, saj so takrat na vrsti odmori.



Slika 8.8: Število zaznanih naprav v predavalnic P1, v četrtek, 2. aprila 2015.

Preden smo signale razvrstili po regijah, smo izvedli še meritve signalov

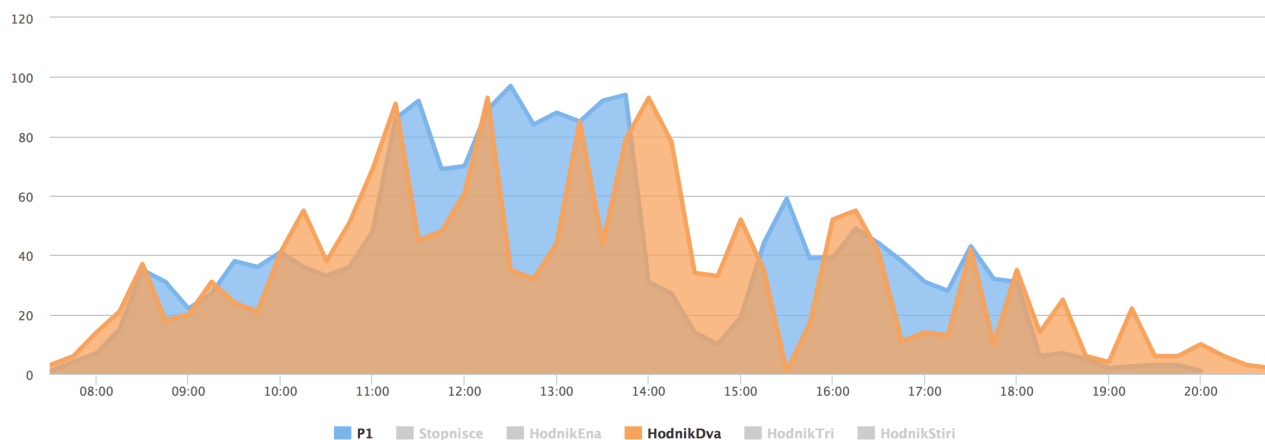
po regijah. V razdelku 7.2.2 smo opisali postopek in pregledali statistiko zajetih podatkov. S pomočjo algoritma za klasifikacijo, ki smo ga definirali v razdelku 6.4, smo klasificirali vse signale in jim določili regije. Nato smo narisali graf z obiskom po regijah. Na sliki 8.8 je prikazan obisk skozi čas za predavalnico P1.



Slika 8.9: Število zaznanih naprav v P1 in zaznan povečan obisk v času predavanj.

Prve obiskovalce predavalnice P1 je algoritem zaznal ob 8.00, sklepamo pa, da se je prvo predavanje začelo ob 8.30, kar so potrdili tudi študentje, ki smo jih povprašali, kdaj začnejo s predavanji ob četrtkih. Ob 9.00 je nastopil prvi odmor in ob 11.00 so se zaključila predavanja. Med 11.15 in 13.45 smo zaznali drugo predavanje. Nato je očitno sledil manjši obisk, zaradi odmora. Od 15.15 pa lahko ponovno vidimo prisotnost ljudi do 18.00. Rezultati se torej zelo dobro ujemajo z urnikom. Zaznana predavanja smo označili na sliki 8.9. Zapisali smo tudi, kateri letnik je imel ob določeni uri predavanja. V času med 11.15 in 13.45 smo zaznali največji obisk, saj so imeli takrat predavanja na vrsti študentje prvega letnika. V povprečju je sistem v tem času zaznal nekaj manj kot 100 mobilnih naprav.

Preverimo še, kaj se je med tem časom dogajalo na hodniku, natančneje v regiji *HodnikDva*. Rezultat je predstavljen na sliki 8.10. Algoritem je zaznal



Slika 8.10: Primerjava obiska med regijo *P1* in *HodnikDva*.

kratkotrajna obdobja povečanega obiska, in sicer ob: 8.30, 9.15, 10.15, 11.15, 12.15, 13.15, 14.15, 15.00, 16.00, 16.15, 17.15 in 18.00. Te ure se zelo dobro ujemajo z odmori, prav tako absolutno število obiskovalcev med predavalnico in hodnikom.

Torej, s tem primerom smo pokazali, da so rezultati klasifikacije smiselni. Natančnejša evalvacija algoritma za klasifikacijo pa sledi v naslednjem, 9. poglavju.

Poglavje 9

Evalvacija klasifikacije signalov Wi-Fi

Problem klasifikacije signalov Wi-Fi in algoritme, ki rešujejo ta problem, smo natančno definirali v 6. poglavju. V tem poglavju pa evalviramo posamezne algoritme s pomočjo izmerjenih signalov, za katere poznamo izvirno regijo. Te meritve so predstavljene v dodatku C, podrobnejša analiza pa je predstavljena v razdelku 7.2.2. Za vsak algoritem klasificiramo 242 vektorjev¹, nato pa izračunamo natančnost klasifikacije (angl. classification accuracy, CA) z naslednjo formulo:

$$CA [\%] = 100 * \frac{\#pravilnih\ klasifikacij}{\#vseh\ opisov\ signala} \quad (9.1)$$

Poleg natančnosti klasifikacije, preverimo, kam posamezen algoritem uvrsti tiste signale, ki izvirajo iz nedefinirane regije. Ob koncu poglavja podrobneje predstavimo izbiro parametrov za klasifikacijo s strojnim učenjem.

9.1 Naključno razporejanje

Algoritem za naključno razporejanje klasificira signal v pravilno regijo z verjetnostjo $\frac{1}{n}$, kjer n predstavlja število regij. Ker smo definirali 6 regij, bi

¹Vektor predstavlja opis zaznanega signala.

algoritem v povprečju pravilno razvrstil $\frac{1}{6}$ oz. **16.7 %** signalov. Naključno razporejanje predstavlja izhodišče (angl. baseline) za nadaljnje algoritme.

9.2 Metoda najmočnejše dostopne točke

V razdelku 6.2 smo definirali algoritem za klasifikacijo z metodo najmočnejše dostopne točke. Če želimo uporabljati ta algoritem, je najprej potrebno definirati preslikavo med dostopnimi točkami in regijami. Za naš praktični preizkus smo definirali naslednjo preslikavo:

najmočnejša DT		regija
DT_{P1}	\implies	P1
-	\implies	HodnikEna
-	\implies	HodnikDva
DT_{SSFRI}	\implies	HodnikTri
DT_{Garaza}	\implies	HodnikStiri
$DT_{Referat}$	\implies	Stopnisce

Na voljo smo imeli le štiri dostopne točke za šest regij in posledično preslikava ne more biti bijektivna. Algoritem tako nikoli ne razvrsti signalov v regiji *HodnikEna* in *HodnikDva*. Kljub temu pa natančnost klasifikacije znaša **57,0 %**, kar je boljše od naključnega razporejanja.

Če izmed 242 vektorjev izločimo vse tiste, ki so bili zajeti v regijah *HodnikEna* in *HodnikDva*, potem algoritem pravilno napove regijo kar v **82,6 %** primerih. To pomeni, da algoritem dela zelo dobro, če imamo na voljo dovolj dostopnih točk in jih lahko namestimo v poljubno regijo. V praksi to seveda predstavlja višje stroške in višjo kompleksnost vzpostavitve sistema, vendar pa v bodoče lahko upravičeno pričakujemo nižje cene dostopnih točk² in s tem nižjo ceno celotnega sistema.

²V pripravi so projekti, ki obljublajo nižje cene dostopnih točk, npr. projekt [14] obljublja ceno pod 10 EUR za vsako dostopno točko.

V primeru izvora signala izven definiranih regij, bi algoritem po trenutni definiciji signal klasificiral v eno izmed regij. Problem bi zopet lahko rešili z dodatnimi dostopnimi točkami, ki bi jih postavili v neznane regije, preslikavo pa bi definirali tako, da te dodatne točke preslikajo v neznano regijo.

9.3 Trilateracija

V razdelku 6.3 smo definirali algoritem za klasifikacijo signalov s trilateracijo. Za potrebe evalvacije algoritma, smo implementirali algoritem trilateracije tako, kot je predstavljen v razdelku 3.6. Ker ta metoda predvideva točno 3 dostopne točke, smo ob implementaciji poskrbeli še za primere, ko signale zaznajo 1, 2 ali 4 dostopne točke:

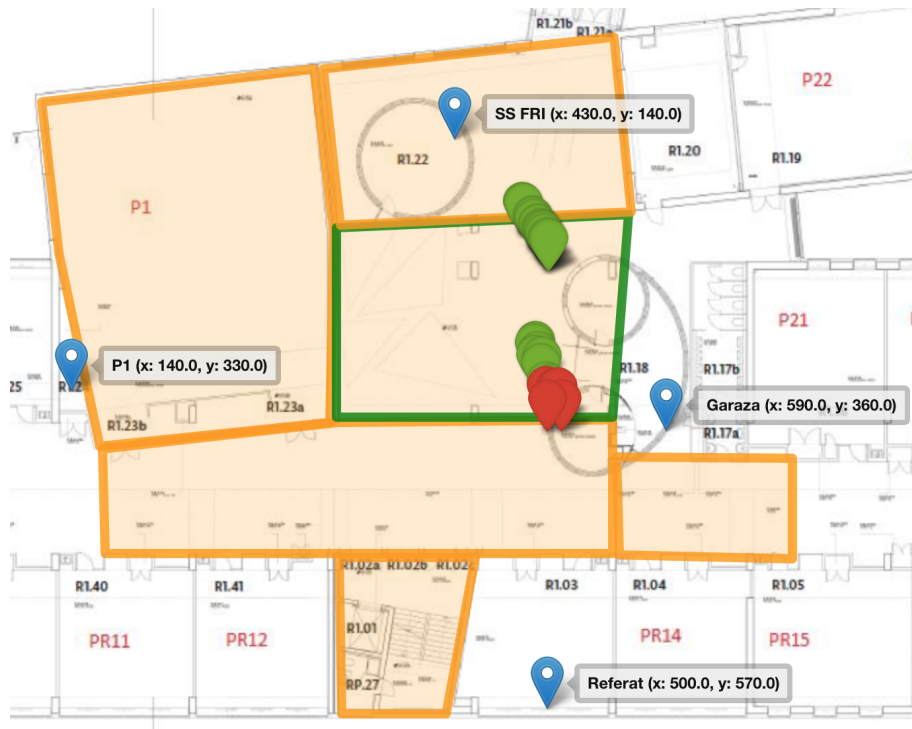
- 4 dostopne točke:** lokacijo smo dobili tako, da smo za vse štiri kombinacije treh dostopnih točk izračunali lokacijo s trilateracijo, nato pa smo za te štiri izračune določili težišče;
- 2 dostopni točki:** postopek trilateracije v tem primeru ni možen, zato smo vrnili le težišče na premici, ki povezuje obe dostopni točki;
- 1 dostopna točka:** vrnili smo lokacijo dostopne točke, nato pa smo določili regijo na podlagi preslikave, ki smo jo določili pri klasifikaciji z metodo najmočnejše dostopne točke.

Izračunali smo natančnost klasifikacije tega algoritma in le v dobri polovici, natančneje **50,4 %**, je algoritem signal uvrstil v pravo regijo. Rezultat je precej slab, še slabši od klasifikacije z metodo najmočnejše dostopne točke.

Očitno ta metoda ni najprimernejša. Zahteva več dela, saj je treba vpeti koordinatni sistem in za vsako regijo definirati poligon. Eden izmed glavnih problemov je, da signal pogosto zaznata le dve oz. ena dostopna točka³. Trilateracija v 2D prostoru pa dela dobro šele takrat, ko signal zaznajo vsaj 3 dostopne točke. Velika pomanjkljivost tega algoritma je, da

³V razdelku 7.2.2 smo ugotovili, da ima le dobra četrtina vektorjev 3 oz. 4 komponente

ne upošteva razporeditve sten. Čeprav je enostaven za vzpostavitev, je na drugi strani določanje lokacije premalo natančno, posledično pa je natančnost klasifikacije zelo slaba.



Slika 9.1: Klasifikacija zajetih signalov v regiji *HodnikDva* - zelena zastavica označuje pravilno klasifikacijo, rdeča pa napačno.

Na drugi strani pa algoritem omogoča razvrščanje signalov iz nedefiniranih regij v neznano regijo, kar npr. metoda najmočnejše dostopne točke ne omogoča.

9.4 Klasifikacija s strojnim učenjem

Strojno učenje potrebuje testne podatke, na podlagi katerih algoritem zgradi model. Zato bomo za ocenjevanje natančnosti uporabili navzkrižno preverjanje (angl. cross validation), sestavljeno iz petih iteracij. Vhodni podatki se razdelijo na 5 enakih delov. Pri vsaki iteraciji se iz $\frac{4}{5}$ podatkov zgradi učna

množica, na kateri klasifikator zgradi model, ta pa se nato testira na preostali $\frac{1}{5}$ podatkov. Končno klasifikacijsko natančnost predstavlja povprečje natančnosti vseh petih iteracij. Na ta način zagotovimo realno primerjavo s preostalimi algoritmi.

Pripravili smo 3 različice vhodnih podatkov, ki se razlikujejo po opisu znanega signala in privzeti vrednosti. Slednja se uporabi takrat, ko dostopna točka ne zazna signala.

1. **moč signala** - osnovni podatki so bili izraženi z močjo signala [dBm]. Za privzeto vrednosti smo izbrali -100 dBm. Ta vrednost predstavlja zelo šibke signale, katere je zelo težko zaznati;
2. **razdalja** - moč signala smo pretvorili v razdaljo [m] s pomočjo modela za pojenjanje signala. Uporabili smo formulo 3.10 s parametrom $n = 5$. Namesto manjkajočih vrednosti, smo vpisali 50 m.
3. **kvaliteta signala** - moč signala smo pretvorili v kvaliteto signala [%] po formuli 3.11. Manjkajoče vrednosti smo nadomestili s kvaliteto 0,0 %.

Uporabili smo štiri različne klasifikatorje: SVM, naivnega Bayesa, logistično regresijo in k najbližjih sosedov. Za vsakega smo izračunali klasifikacijsko natančnost (CA), za vse 3 različice vhodnih podatkov. Testiranja smo izvedli s pomočjo knjižnice *scikit-learn*. Rezultati so prikazani v tabeli 9.1.

klasifikator	moč signala	razdalja	kvaliteta signala
SVM	0,8187	0,8433	0,8512
naivni Bayes	0,7730	0,7855	0,8349
logistična regresija	0,7895	0,7730	0,7895
k najbližjih sosedov	0,8349	0,8348	0,7730

Tabela 9.1: Primerjava klasifikacijske natančnosti, glede na klasifikator in tip vhodnih podatkov.

Najboljše se je odrezal SVM na podatkih, ki smo jih opisali s kvaliteto signala. Klasifikacijska natančnost znaša **85,1 %**, kar je daleč najboljše izmed vseh predstavljenih metod. V nadaljevanju bomo vse podatke predstavili na podlagi klasifikatorja SVM in podatkov predstavljenih s kvaliteto signala.

S strojnim učenjem lahko tudi ugotovimo, katere regije so si med seboj najbolj podobne. Na ta način lahko preverimo, kako se obnese klasifikacija, če definiramo regije po stenah in kako, če ne. Tabela 9.2 predstavlja matriko zamenjav (angl. confusion matrix). Na levi so zapisane dejanske regije, zgoraj pa napovedan razred s strani klasifikatorja. Vidimo, da klasifikator regijo *P1* vedno razvrsti pravilno. Zelo dobro se odreže tudi v primeru regij *HodnikTri*, *HodnikStiri* in *Stopnisce*. Največ težav pa ima pri razlikovanju med regijama *HodnikEna* in *HodnikDva*, predvsem slednjo zelo pogosto zameša z regijo *HodnikTri*. Torej, največ je težav pri razlikovanju med regijami *HodnikEna*, *HodnikDva* in *HodnikTri*. To so pa ravno tiste regije, ki niso ločene med seboj s stenami. Kljub temu pa klasifikator še vedno z več kot 50 % uspešnostjo loči med temi regijami.

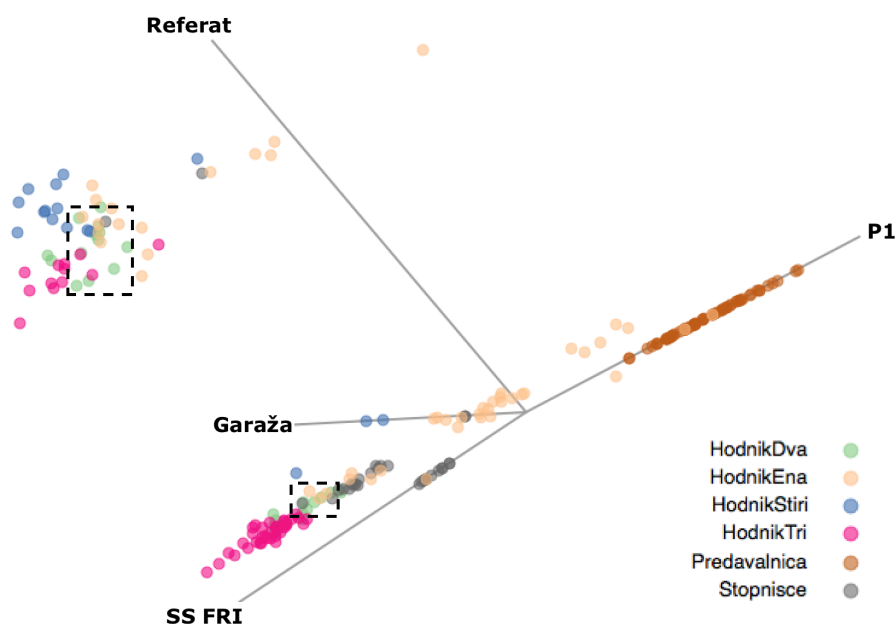
Regija	Hodnik2 ⁴	Hodnik1	Hodnik4	Hodnik3	P1	Stopnisce	#
Hodnik2	50.0	7.7	0.0	30.8	0.0	11.5	26
Hodnik1	12.2	65.3	0.0	0.0	6.1	16.3	49
Hodnik4	0.0	6.2	93.8	0.0	0.0	0.0	16
Hodnik3	4.2	0.0	0.0	93.8	0.0	2.1	48
P1	0.0	0.0	0.0	0.0	100	0.0	73
Stopnisce	6.7	6.7	0.0	0.0	0.0	86.7	30
#	23	37	15	53	76	38	242

Tabela 9.2: Matrika zamenjav: na levi so dejanske regije, zgoraj pa napovedane.

Za boljšo predstavbo, smo s pomočjo metode glavnih komponent (angl. principal component analysis, PCA) vizualizirali podatke. Rezultat je predstavljen na sliki 9.2. Uporabili smo podatke, predstavljene s kvaliteto si-

⁴Imena regij smo zapisali skrajšano. Namesto *HodnikDva*, *Hodnik2*.

gnala. Lepo se vidi, da se večino podatkov lahko zelo enostavno klasificira. S črtkastima kvadratoma smo označili dve območji, kjer se podatki iz različnih regij prekrivajo med seboj. V kvadratih se nahajajo podatki iz regij *HodnikEna*, *HodnikDva*, *HodnikTri* in *HodnikStiri*. To so pa ravno tiste regije, ki niso v celoti omejene s stenami. Še en dokaz več, da je pametno regije postaviti po stenah in s tem olajšati delo klasifikatorju.



Slika 9.2: Vizualizacija podatkov s pomočjo metode glavnih komponent (angl. Principal Component Analysis, PCA).

Metodo smo prilagodili tako, da smo upoštevali tudi neznane regije. Knjižnica *scikit-learn* ima funkcijo *predict_proba*, ki vrne verjetnost klasifikacije za vsako regijo. Definirali smo prag, ki določa, kako prepričan mora biti klasifikator, da nek signal dejansko razvrsti v eno izmed regij. Prag smo določili pri 70 %. S tem smo signale, ki prihajajo iz nedefiniranih regij, razvrstili v *neznano regijo*.

Poglavje 10

Sklepne ugotovitve

V zadnjih letih se hitro širijo brezžična omrežja Wi-Fi, vse več pa je tudi pametnih mobilnih naprav, ki se lahko na ta omrežja priključijo. Medtem ko mobilne naprave iščejo brezžična omrežja, ves čas oddajajo signale, katere je možno zaznati z dostopnimi točkami. Na ta način lahko preštejemo, koliko mobilnih naprav se nahaja v bližini dostopne točke. Ker je mobilnih naprav vse več, lahko zelo dobro ocenimo relativni obisk skozi čas.

Razvili smo poceni sistem za zajem signalov Wi-Fi, ki bazira na računalnikih Raspberry Pi. Z njim smo v času med 13. marcem in 12. majem 2015 zajemali podatke na Fakulteti za računalništvo in informatiko v Ljubljani. Nato smo podatke natančno analizirali. Ugotovili smo, da lahko že z eno dostopno točko zaznamo relativni obisk skozi čas. Izmerjen obisk smo podrobneje primerjali z urnikom Fakultete in rezultati so se zelo dobro ujemali. Pri tej analizi smo upoštevali vse zaznane signale iz okolice dostopne točke. V naslednji fazi pa smo poizkušali določiti še izvor signala, do prostora oz. regije natančno.

Preverili smo, kateri produkti obstajajo na trgu, ki omogočajo lociranje naprav, na podlagi signala Wi-Fi. Rešitve ponujajo več ali manj vsi večji proizvajalci opreme Wi-Fi. Primerjali smo Cisco MSE, Meraki Presense Analytics in Ruckus SPoT ter ugotovili, da je cena teh sistemov zelo visoka, saj proizvajalci pogojujejo uporabo z dragimi dostopnimi točkami. Ker so pro-

dukti prvenstveno namenjeni natančnemu določanju lokacije, zahtevajo tudi časovno potratno kalibracijo.

Poleg produktov, smo poiskali tudi najboljše metode za določanje lokacije v notranjih prostorih. Podrobneje smo predstavili metodo RADAR, ki obljublja natančnost 2-3 metre. Glavni problem te metode je, da je potrebno opraviti natančne meritve signala po prostorih, kar zahteva veliko časa. Opisali smo še metodo trilateracije, za katero zamudna kalibracija ni potrebna, vendar pa je njena natančnost veliko slabša, nekje med 5-10 metrov.

Nobena izmed obstoječih metod ni idealna, zato smo sklenili kompromis. Ker želimo najti le izvorno regijo signala, ni potrebe po določanju natančne lokacije. Problem smo tako prevedli na klasifikacijo signalov po regijah. Razvili smo tri različne metode: klasifikacijo z metodo najmočnejše dostopne točke, klasifikacijo s trilateracijo in klasifikacijo s strojnim učenjem. Na Fakulteti smo postavili sistem s štirimi dostopnimi točkami, območje merjenja pa smo razdelili na šest različnih regij - dve od teh smo definirali točno po stenah, štiri pa le delno. Nato smo natančno premerili, katera izmed metod se obnese najboljše. Z metodo najmočnejše dostopne točke, smo dosegli klasifikacijsko natančnost 57,0 %, s trilateracijo 50,4 %, s strojnim učenjem pa 85,1 %.

Klasifikacija s strojnim učenjem poleg dobre klasifikacijske natančnosti, prinaša tudi druge prednosti v primerjavi z ostalimi metodami. Meritve signalov so zelo enostavne, saj za vsako meritev ni potrebe po vnašanju koordinat, ampak zadostuje le ime regije. Ne rabimo zemljevidov z razporeditvijo prostorov, prav tako ni potrebe po vpeljavi koordinatnega sistema. Metoda se obnese najboljše, če regije predstavljajo prostore, ločene s stenami. V tem primeru klasifikacijska natančnost znaša prek 93 %. Pri razvoju te metode, smo testirali različne klasifikatorje, najboljše pa se je obnesel SVM.

Poleg signalov Wi-Fi, smo zajemali tudi vremensko napoved, s pomočjo katere smo iskali korelacijo med vremenom in obiskom Fakultete. Izkaže se, da na obisk zares vpliva tudi vreme, vendar pa ni edini dejavnik. Poleg vremena na obisk zelo močno vpliva dan v tednu. Med delovnim tednom smo

v povprečju zaznali najmanjši obisk ob petkih, največjega pa ob četrtnih in ponedeljkah. Napoved obiska samo na podlagi vremena ni možna, lahko pa s kombinacijo vremenske napovedi in dneva v tednu, precej dobro napovemo obisk za določen dan. Splošne rešitve, na podlagi katere bi lahko napovedovali obisk, nismo našli.

Pri zajemu signalov Wi-Fi, je potrebno zajeti tudi naslove MAC, na podlagi katerih ločimo posamezne mobilne naprave med seboj. Ti podatki so s stališča zasebnosti zelo občutljivi. Zato smo v okviru magistrske naloge pripravili rešitev, ki poskuša zagotoviti visok nivo zasebnosti. Uporabili smo zgoščevalno funkcijo *SHA256*, ki zgosti naslov MAC, skupaj z datumom zajema in lokacijo. Na ta način preprečimo vpogled, kako pogosto se mobilna naprava vrača, prav tako pa ni možno določiti povezave z ostalimi lokacijami.

S to magistrsko nalogo smo dokazali, da je možno s poceni opremo ter klasifikacijo signalov Wi-Fi s strojnim učenjem, zelo natančno določiti regijo izvora signalov Wi-Fi. To nam omogoča, da lahko na enostaven način merimo obisk po različnih regijah. Primerjamo lahko obisk različnih predavanj na konferencah, obisk različnih trgovin v trgovskih središčih, odkrijemo lahko, v katerem prostoru smo izgubili svoj mobilni telefon ali tablico, možno je tudi napovedovati, kdaj lahko pričakujemo gnečo na blagajnah v velikih trgovinah, itd.

10.1 Nadaljnje delo

V okviru magistrske naloge smo razvili novo metodo za klasifikacijo signalov Wi-Fi. Preizkusili smo štiri klasifikacijske algoritme (SVM, naivni Bayes, logistična regresija in k najbližjih sosedov) in tri različne tipe vhodnih podatkov. Verjamemo, da je tu še nekaj prostora za napredek. Morda obstaja boljša kombinacija klasifikatorja, vhodnih podatkov in parametrov. Metodo smo preizkusili le na dveh lokacijah, dobro pa bi bilo, da se preizkusi še v drugačnih okoliščinah, npr. trgovskih središčih, kinih, in še kje druge.

Ker klasifikacija ni 100 % natančna, se včasih zgodi, da kakšna na-

prava “skače” med dvema regijama. Predvidevamo, da bi lahko ta problem rešili z medianinim filtrom, ki bi povprečil napovedi regij skozi čas. Zaradi obsežnosti magistrske naloge, filtra nismo preizkusili v praksi, domnevamo pa, da bi lahko na ta način še izboljšali klasifikacijsko natančnost.

Zanimiva bi bila tudi študija, ki bi dala odgovor, koliko dostopnih točk potrebujemo za določen razpored prostorov in kam se jih najbolj splača postaviti. Želeli bi, da je klasifikacijska natančnost še vedno dobra, hkrati pa da za tako pokritje uporabimo čim manj dostopnih točk.

Literatura

- [1] Is looking for Wi-Fi access points purely passive? <http://superuser.com/questions/128166/is-looking-for-wi-fi-access-points-purely-passive>. Citirano 22.1.2015.
- [2] Microsoft MSDNA: Definition of WLAN signal quality. <https://msdn.microsoft.com/en-us/library/windows/desktop/ms706828%28v=vs.85%29.aspx?f=255&MSPPErr=-2147217396>. Citirano 20.3.2015.
- [3] Razpis konference WebCamp 2015, Ljubljana, marec 2015. <http://2015.webcamp.si/>. Citirano 3.4.2015.
- [4] Seznam podprtih SD kartic za Raspberry. http://elinux.org/RPi_SD_cards. Citirano 3.3.2015.
- [5] Smart cities and communities: The european innovation partnership on smart cities and communities. <http://ec.europa.eu/eip/smartcities/>. Citirano 28.2.2015.
- [6] Tekmovanje ACM iz računalništva in informatike. <http://rtk.ijs.si/razpis.html#znanje>, March 2015. Citirano 29.5.2015.
- [7] IEEE Standards Associaton. OUI: baza MAC naslovov in proizvajalcev mrežnih naprav. <http://standards-oui.ieee.org/oui.txt>. Citirano 18.12.2014.

-
- [8] IEEE Standards Associaton. Standard IEEE 802.11. <http://standards.ieee.org/about/get/802/802.11.html>. Citirano 19.12.2014.
- [9] P. Bahl and V.N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784 vol.2, 2000.
- [10] Paramvir Bahl, Venkata N Padmanabhan, and Anand Balachandran. Enhancements to the RADAR user location and tracking system. Technical report, technical report, Microsoft Research, 2000.
- [11] Marco V. Barbera, Alessandro Epasto, Alessandro Mei, Vasile C. Perta, and Julinda Stefa. Signals from the crowd: Uncovering social relationships through smartphone probes. In *Proceedings of the 2013 Conference on Internet Measurement Conference, IMC '13*, pages 265–276, New York, NY, USA, 2013. ACM.
- [12] Misra Bhupinder. Analyze of iOS MAC randomization. <http://blog.airtightnetworks.com/ios8-mac-randomization-analyzed/>, September 2014. Citirano 17.4.2015.
- [13] Inc. Cisco Systems. White paper: Meraki presence analytics. Technical report, May 2013.
- [14] Next Thing Co. C.H.I.P. - the world's first nine dollar computer. <https://www.kickstarter.com/projects/1598272670/chip-the-worlds-first-9-computer>, May 2015. Citirano 27.5.2015.
- [15] B Cook, G Buckberry, I Scowcroft, J Mitchell, and T Allen. Indoor location using trilateration characteristics. In *Proc. London Communications Symposium*, pages 147–150, 2005.

-
- [16] George Dedes and Andrew G Dempster. Indoor GPS positioning. In *Proceedings of the IEEE Semiannual Vehicular Technology Conference*. Citeseer, 2005.
 - [17] Zahid Farid, Rosdiadee Nordin, and Mahamod Ismail. Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications*, 2013, 2013.
 - [18] Dennis Fisher. Research finds MAC address hashing not a fix for privacy problems. <https://threatpost.com/research-finds-mac-address-hashing-not-a-fix-for-privacy-problems/104893>, March 2014. Citirano 12.6.2015.
 - [19] James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes. *Computer Graphics: Principles and Practice (2nd Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.
 - [20] Cisco Internet Business Solution Group. How Wi-Fi will change the mobile industry as we know it. 2012.
 - [21] Guangjie Han, Deokjai Choi, and Wontaek Lim. Reference node placement and selection algorithm based on trilateration for indoor sensor networks. *Wireless Communications and Mobile Computing*, 9(8):1017–1027, 2009.
 - [22] Todd D Hodes, Randy H Katz, Edouard Servan-Schreiber, and Lawrence Rowe. Composable ad-hoc mobile services for universal interaction. In *Proceedings of the 3rd annual ACM/IEEE international conference on Mobile computing and networking*, pages 1–12. ACM, 1997.
 - [23] Huawei. LTE small cell vs. WiFi user experience. 2013.
 - [24] Stefan Kraxberger, Günther Lackner, and Udo Payer. WLAN location determination without active client collaboration. In *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*, pages 1188–1192. ACM, 2010.

-
- [25] James Kurose. *Computer networking: a top-down approach*. Addison-Wesley, Boston, 2010.
- [26] Le T Nguyen and Joy Zhang. Wi-Fi fingerprinting through active learning using smartphones. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 969–976. ACM, 2013.
- [27] Philip Nobles, Shahid Ali, and Howard Chivers. Improved estimation of trilateration distances for indoor wireless intrusion detection. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1):93–102, 2011.
- [28] National Research Council (U.S.). Committee on the Future of the Global Positioning System, National Academy of Public Administration, Commission on Engineering, and Technical Systems. *The Global Positioning System: A Shared National Asset*. Romantic Reassessment; 118. National Academies Press, 1995.
- [29] World Weather Online. Dokumentacija programskega vmesnika World Weather Online. <http://www.worldweatheronline.com/api/docs/>. Citirano 17.4.2015.
- [30] Foster Provost and Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-analytic Thinking*. O'Reilly Media, Inc., 1st edition, 2013.
- [31] Purplemath. Box-and-Whisker plots: Quartiles, boxes, and whiskers. <http://www.purplemath.com/modules/boxwhisk.htm>. Citirano 30.5.2015.
- [32] Radiotap.org. Podrobnejši opis glave radiotap. <http://www.radiotap.org/defined-fields>. Citirano 6.1.2015.
- [33] Anmol Sheth, Sergiu Nedevschi, Rabin Patra, Sonesh Surana, Eric Brewer, and Lakshminarayanan Subramanian. Packet loss characterization

- in WiFi-based long distance networks. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 312–320. IEEE, 2007.
- [34] Telekom Slovenije. WiFree: Brezplačno Wi-Fi brezžično omrežje v Ljubljani. <http://www.wifreeljubljana.si/>. Citirano 16.12.2014.
- [35] Asim Smailagic and David Kogan. Location sensing and privacy in a context-aware computing environment. *Wireless Communications, IEEE*, 9(5):10–17, 2002.
- [36] William Stallings. *Network security essentials: applications and standards*. Pearson, Boston, 2014.
- [37] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems (TOIS)*, 10(1):91–102, 1992.
- [38] Wireshark. Dokumentacija orodja tshark. <https://www.wireshark.org/docs/man-pages/tshark.html>. Citirano 12.3.2015.
- [39] Oliver Woodman and Robert Harle. RF-based initialisation for inertial pedestrian tracking. *Pervasive Computing*, pages 238–255, 2009.
- [40] Rong-Hou Wu, Yang-Han Lee, Hsien-Wei Tseng, Yih-Guang Jan, and Ming-Hsueh Chuang. Study of characteristics of RSSI signal. In *Industrial Technology, 2008. ICIT 2008. IEEE International Conference on*, pages 1–3. IEEE, 2008.
- [41] Fakulteta za računalništvo in informatiko. Novica o zaključku tekmovanja RTK. http://www.fri.uni-lj.si/si/novice_in_dogodki/aktivne/19359/novica.html. Citirano 22.5.2015.
- [42] Fakulteta za računalništvo in informatiko. Urnik Fakultete za računalništvo in informatiko. <https://urnik.fri.uni-lj.si>. Citirano 1.4.2015.

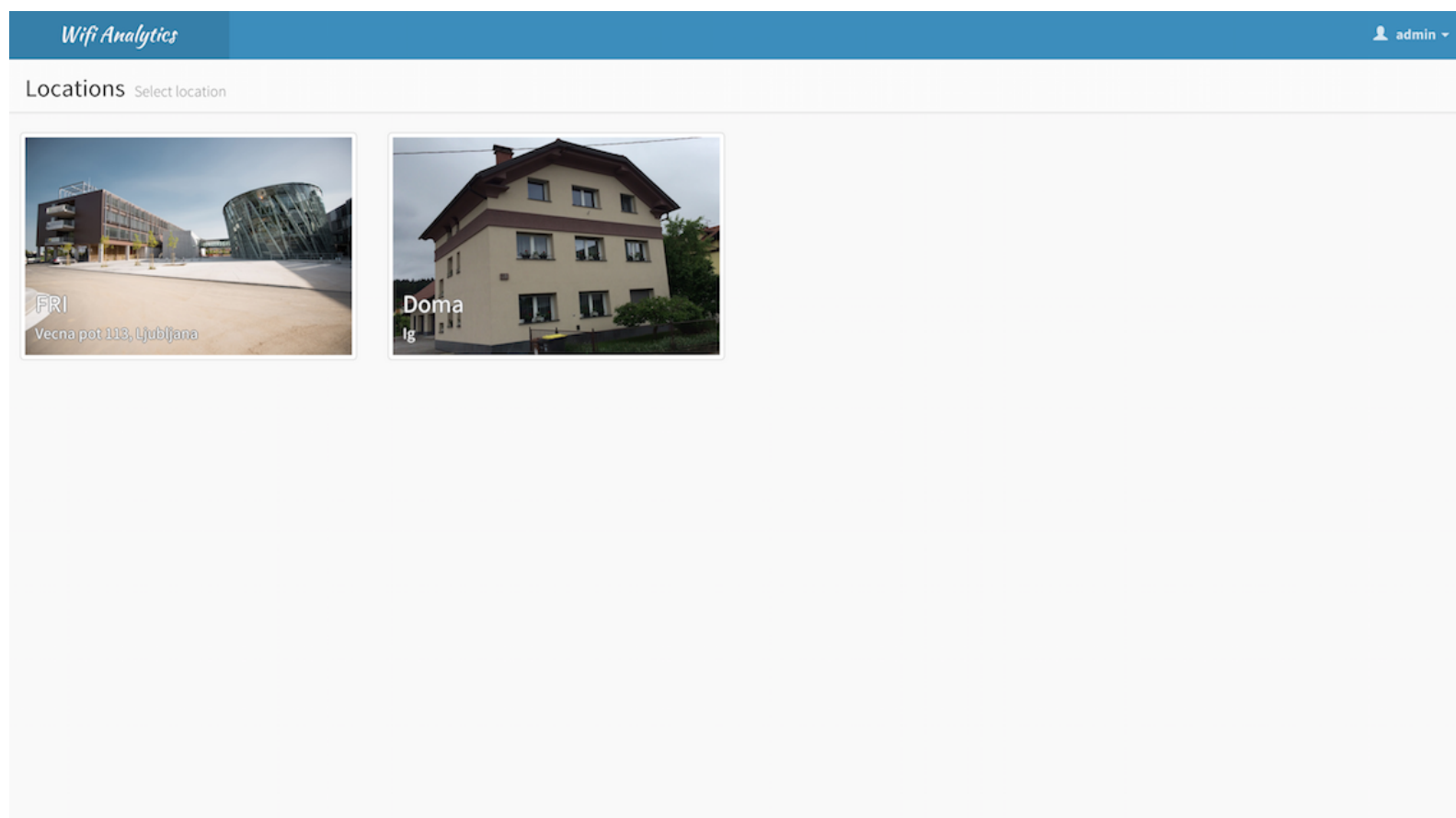
Dodatek A

Spletni portal *Wi-Fi Analytics*

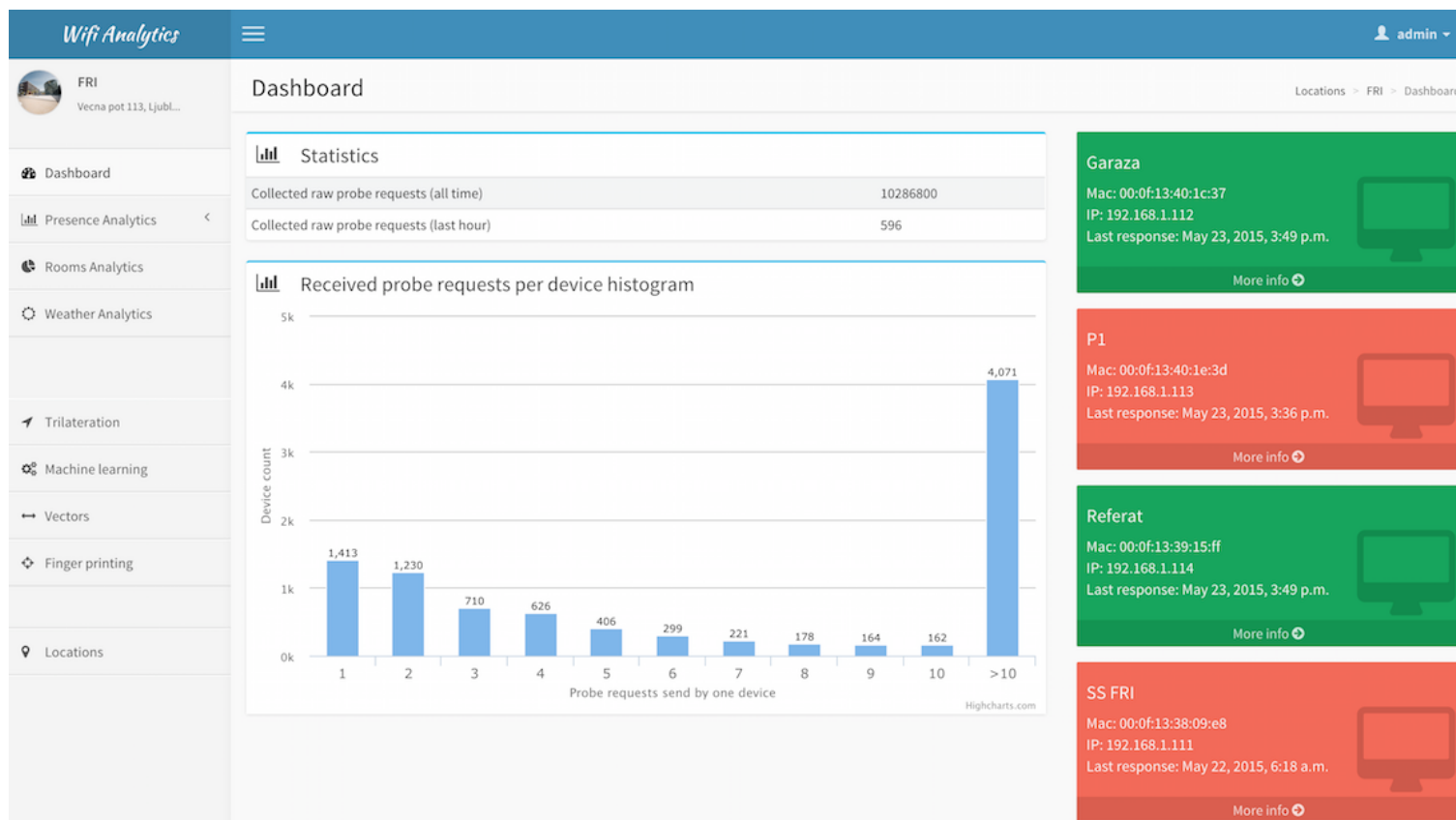
V okviru magistrske naloge smo razvili tudi spletni portal, ki smo ga poimenovali *Wi-Fi Analytics*. V prvi vrsti je namenjen predstavitvi rezultatov te magistrske naloge, hkrati pa služi tudi kot ogrodje za testiranje algoritmov za določanje lokacije. Portal je narejen s sodobnimi spletnimi tehnologijami: HTML5, CSS3, jQuery, Twitter Bootstrap, Django, Highcharts, Scikit Learn, Matplotlib, itd. Portal črpa podatke neposredno iz sistema za zajem podatkov Wi-Fi preko vmesnika REST in ponuja naslednje funkcionalnosti:

- preklapljanje med različni lokacijami;
- spremljanje stanja dostopnih točk;
- prikaz statistike zajetih podatkov;
- analitiko zajetih podatkov za eno dostopno točko;
- vnašanje regij;
- merjenje signalov znotraj posamezne regije;
- iskanje korelacije med vremensko napovedjo in zajetimi podatki;
- testiranje algoritma za trilateracijo.

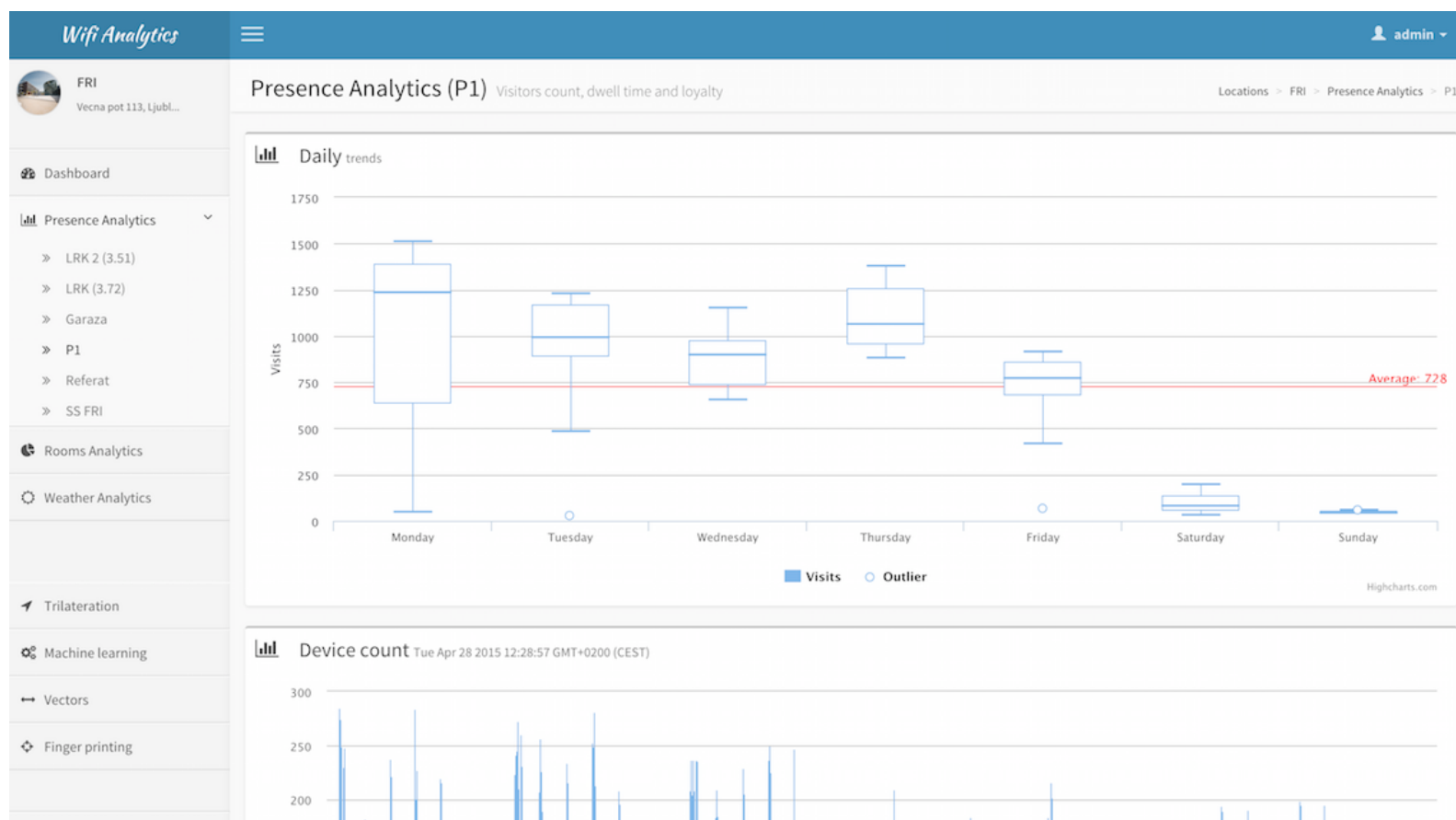
V nadaljevanju sledi predstavitev grafičnega vmesnika in kratek opis, čemu je namenjen posamezni pogled.



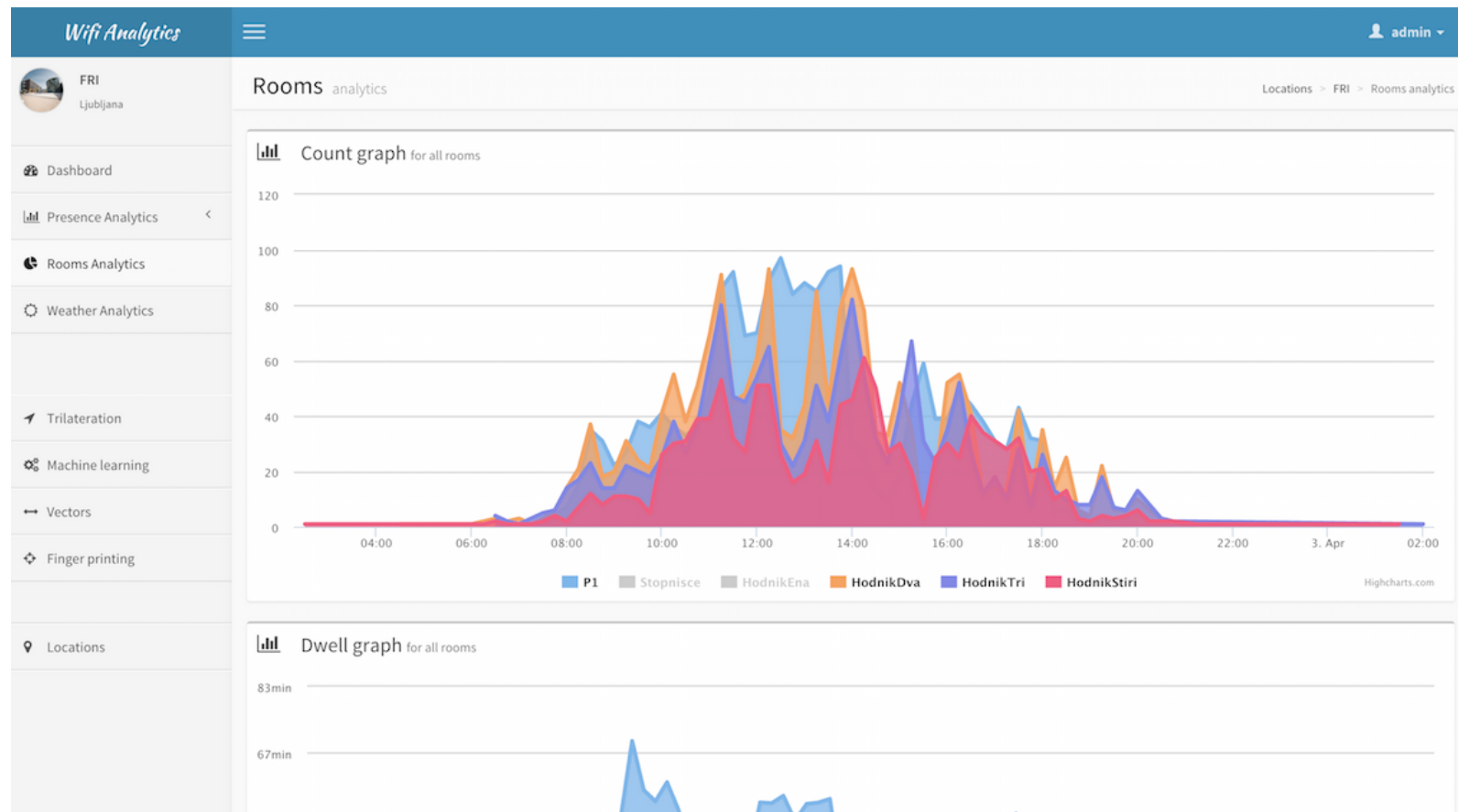
Slika A.1: Lokacije - uporabnik lahko preklaplja med različnimi lokacijami. Uporabnik vidi le tiste lokacije, za katere ima dovolj pravic.



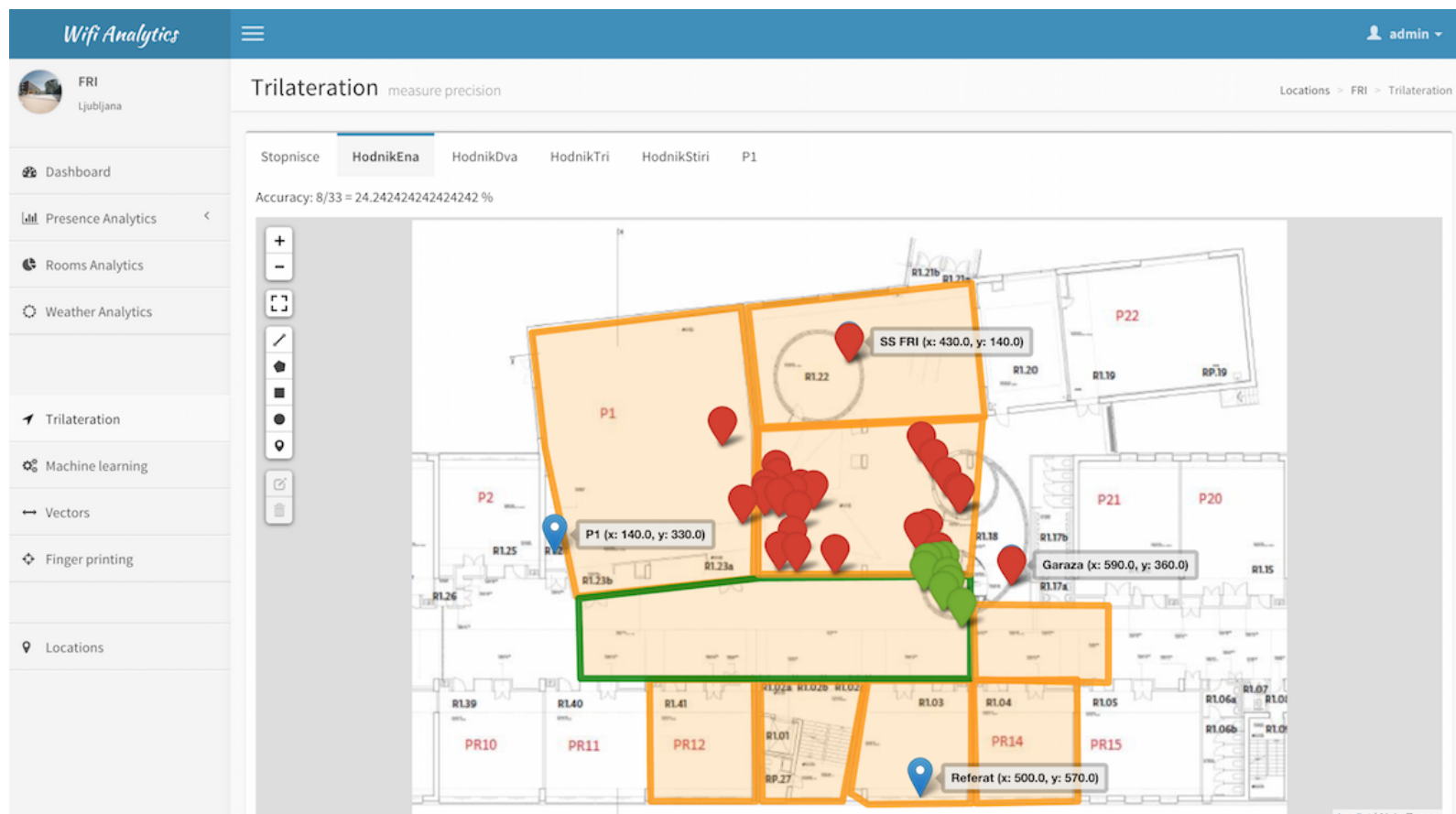
Slika A.2: *Dashboard* služi za predstavitev statistike zajetih podatkov in trenutnega stanja dostopnih točk. Za vsako dostopno točko je prikazan pravokotnik, ki je obarvan z zeleno oz. rdečo barvo. Zelena barva predstavlja točke, ki so poslale podatke v zadnjih 5 minutah, rdeča pa tiste, ki v zadnjih 5 minutah niso poslale podatkov.



Slika A.3: *Presence Analytics* - ta pogled prikazuje zajete podatke za vsako dostopno točko posebej in sicer: število obiskov po dnevih v tednu, število obiskov skozi čas in povprečen čas zadrževanja v bližini dostopne točke.



Slika A.4: *Rooms Analytics*: pogled, ki prikazuje obiskanost posameznih regij in koliko časa so se obiskovalci zadrževali v različnih regijah v povprečju. Na sliki so prikazani rezultati za 2. april.



Slika A.5: Trilateracija: ta pogled je namenjen testiranju, kako dobro dela algoritem trilateracije. Na sliki so prikazane vse lokacije signalov, ki so bile zajete v regiji *HodnikEna*. Zelena barva predstavlja pravilno lociranje, rdeča pa napačno. Vidimo lahko, da je bila trilateracija v tem primeru uspešna le 24 %.

Wifi Analytics

admin

FRI

Vecna pot 113, Ljubljana

Dashboard

Presence Analytics

Rooms Analytics

Weather Analytics

Trilateration

Machine learning

Vectors

Finger printing

Locations

Finger printing

Locations > FRI > Finger printing

New interval

Mac address

54:ae:27:69:58:41

Room

Stopnische P1 PR14 PR12 Referat HodnikEna HodnikDva HodnikTri HodnikStiri

Interval recording

Start

Finger printing intervals

Stopnische

	Device MAC	#raw probe requests	Time start	Time stop	Delta
1	54:ae:27:69:58:41	213	March 26, 2015, 6:04 p.m.	March 26, 2015, 6:04 p.m.	0:00:39.088000
2	54:ae:27:69:58:41	361	April 13, 2015, 6:34 p.m.	April 13, 2015, 6:36 p.m.	0:01:41.163000

P1

	Device MAC	#raw probe requests	Time start	Time stop	Delta
1	54:ae:27:69:58:41	400	March 26, 2015, 6:13 p.m.	March 26, 2015, 6:15 p.m.	0:01:52.086000
2	54:ae:27:69:58:41	39	April 13, 2015, 5:58 p.m.	April 13, 2015, 6 p.m.	0:02:25.734000
3	54:ae:27:69:58:41	390	April 13, 2015, 6:02 p.m.	April 13, 2015, 6:05 p.m.	0:03:00.470000
4	54:ae:27:69:58:41	0	April 13, 2015, 6:06 p.m.	April 13, 2015, 6:06 p.m.	0:00:05.049000

Slika A.6: *Finger printing* omogoča zajem podatkov v znanih regijah. Ko se uporabnik postavi v regijo, vpiše svoj naslov MAC, izbere trenutno regijo in pritisne *Start*. Zaželeno je, da se uporabnik med tem časom premika po prostoru in spodbudi mobilno napravo, da pošlje čim več signalov (npr. s prižiganjem/ugašanjem vmesnika Wi-Fi). Po nekaj minutah se merjenje signala zaključi s pritiskom na gumb *Stop*.

109

The screenshot displays the 'Wifi Analytics - Control Panel' interface. At the top, there is a dark header with the title 'Wifi Analytics - Control Panel' on the left and the user 'admin' on the right. Below the header is a light gray bar with a 'Home' link. The main content area is titled 'Site administration' and is divided into two columns.

Left Column: Site administration

- Analytics**
 - Daily analyticss: + Add ≡ Change
 - Paths: + Add ≡ Change
 - Presences: + Add ≡ Change
 - Rooms analyticss: + Add ≡ Change
 - Vectors: + Add ≡ Change
- Authentication and Authorization**
 - Groups: + Add ≡ Change
 - Users: + Add ≡ Change
- Collector**
 - Finger printing intervals: + Add ≡ Change
 - Locations: + Add ≡ Change
 - Places: + Add ≡ Change
 - Probe request collectors: + Add ≡ Change
 - Raw probe requests: + Add ≡ Change
 - Rooms: + Add ≡ Change
 - Weathers: + Add ≡ Change
- Graphs**
 - Device counts: + Add ≡ Change
 - Dwell times: + Add ≡ Change
 - Loyaltys: + Add ≡ Change
 - Probe requests histograms: + Add ≡ Change
- Web**
 - Device count graph points: + Add ≡ Change

Right Column: Recent Actions

My Actions

- ≡ <Location (name: Doma, address: Ig)> Location
- ≡ <Location (name: Doma, address: Ig)> Location
- ≡ <Location (name: FRI, address: Vecna pot 113, Ljubljana)> Location
- × <Room (name: P22, location: <Location (name: FRI, address: Vecna pot 113, Ljubljana)>)> Room
- + demo User
- ≡ <Room (name: HodnikStiri, location: <Location (name: FRI, address: Vecna pot 113, Ljubljana)>)> Room
- ≡ <Room (name: HodnikTri, location: <Location (name: FRI, address: Vecna pot 113, Ljubljana)>)> Room
- ≡ <Room (name: HodnikDva, location: <Location (name: FRI, address: Vecna pot 113, Ljubljana)>)> Room

Slika A.7: Nadzorna plošča omogoča nastavljanje parametrov in spreminjanje vrednosti objektov v bazi z operacijami *CRUD* (*create, read, update, delete*). Podatke lahko ureja samo administrator, ki lahko dodaja tudi nove uporabnike. Za vsakega uporabnika mora specificirati tudi katere lokacije lahko vidi.

Dodatek B

Vreme in obisk po dnevih

V tabeli so predstavljeni podatki za obdobje med 13.3. in 12.5.2015. Za vsak dan je prikazano število zaznanih intervalov prisotnosti za vsako dostopno točko posebej in vremenski podatki.

Datum	DT_{SSFRI}	DT_{Garaza}	DT_{P1}	$DT_{Referat}$	Temp.	Opis vremena
13.3.	1108	0	0	766	11	Partly Cloudy
14.3.	1793	0	0	1339	10	Sunny
15.3.	8	0	0	8	10	Overcast
16.3.	2321	2012	1240	1537	12	Cloudy
17.3.	2189	2479	1189	1345	12	Sunny
18.3.	2365	2421	1153	1219	12	Partly Cloudy
19.3.	2553	2517	1376	1262	14	Sunny
20.3.	1476	1511	874	781	15	Partly Cloudy
21.3.	400	400	194	190	13	Sunny
22.3.	48	58	46	3	14	Overcast
23.3.	2215	2363	1512	1091	12	Sunny
24.3.	2018	2083	1228	1055	16	Partly Cloudy
25.3.	1959	1863	963	895	11	Light rain
26.3.	2306	2336	1299	1109	17	Light rain shower
27.3.	1526	1537	746	760	9	Patchy light drizzle
28.3.	113	105	85	84	16	Sunny
29.3.	59	51	60	67	16	Sunny
30.3.	2264	2189	1463	1137	14	Partly Cloudy
31.3.	2042	1995	1143	997	11	Light drizzle
1.4.	1858	1902	976	965	13	Partly Cloudy

Datum	DT_{SSFRI}	DT_{Garaza}	DT_{P1}	$DT_{Referat}$	Temp.	Opis vremena
2.4.	2396	2266	1127	1215	12	Patchy rain nearby
3.4.	1479	1487	804	759	16	Sunny
4.4.	99	107	80	72	11	Patchy rain nearby
5.4.	47	50	44	50	9	Sunny
6.4.	49	45	45	47	10	Partly Cloudy
7.4.	1948	1919	995	962	10	Sunny
8.4.	1916	2063	832	1012	16	Cloudy
9.4.	2101	2454	880	1319	16	Sunny
10.4.	1446	1611	901	964	20	Sunny
11.4.	155	142	152	86	18	Sunny
12.4.	48	46	46	48	22	Sunny
13.4.	1781	1886	1233	1019	24	Sunny
14.4.	1582	1754	846	846	18	Partly Cloudy
15.4.	1840	2061	702	994	22	Sunny
16.4.	2376	2510	1005	1262	22	Overcast
17.4.	174	182	68	56	13	Light drizzle
18.4.	42	82	31	43	8	Light rain
19.4.	76	115	51	67	13	Sunny
20.4.	1930	2377	1313	1006	21	Sunny
21.4.	1625	1971	934	883	23	Sunny
22.4.	1533	1660	654	696	24	Sunny
23.4.	2022	2023	945	842	20	Sunny
24.4.	1164	1247	658	602	20	Patchy rain nearby
25.4.	61	86	48	41	20	Sunny
26.4.	42	52	42	42	19	Patchy rain nearby
27.4.	90	68	46	48	20	Cloudy
28.4.	230	292	59	112	15	Patchy light rain
29.4.	120	243	60	106	18	Sunny
30.4.	87	181	54	82	18	Patchy light drizzle
1.5.	10	54	46	41	16	Light rain shower
2.5.	12	56	45	50	22	Patchy rain nearby
3.5.	0	63	45	45	21	Light rain shower
4.5.	1617	1843	1021	740	25	Partly Cloudy
5.5.	1363	1741	746	777	27	Sunny
6.5.	1466	1690	770	892	28	Partly Cloudy
7.5.	1858	1998	908	1074	22	Sunny
8.5.	1182	1425	702	788	23	Sunny

Datum	DT_{SSFRI}	DT_{Garaza}	DT_{P1}	$DT_{Referat}$	Temp.	Opis vremena
9.5.	93	107	49	66	25	Sunny
10.5.	46	79	44	49	26	Sunny
11.5.	1637	1918	1168	930	24	Sunny
12.5.	1609	1898	855	895	28	Sunny

Dodatek C

Meritve signalov po regijah

26.3. in 13.4.2015 smo opravili meritve signalov v šestih regijah: *Stopnisce*, *HodnikEna*, *HodnikDva*, *HodnikTri*, *HodnikStiri* in *P1*. Signale smo generirali z mobilnim telefonom iPhone 5S in sicer tako, da smo vsakih 5 sekund ugasnili in prižgali vmesnik Wi-Fi. V vsaki regiji smo se zadržali 3-5 min in med tem časom so dostopne točke zaznale 242 signalov. Vprašaj (?) pomeni, da dostopna točka (DT) ni zaznala signala. Vse vrednosti so podane z *dBm*.

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
1.	26.3.2015	?	?	?	-83.0	Stopnisce
2.	26.3.2015	?	-87.0	?	?	Stopnisce
3.	26.3.2015	?	?	?	-78.6	Stopnisce
4.	26.3.2015	?	?	?	-84.5	Stopnisce
5.	26.3.2015	?	?	?	-84.5	Stopnisce
6.	26.3.2015	?	?	?	-79.5	Stopnisce
7.	26.3.2015	?	-87.0	?	?	Stopnisce
8.	26.3.2015	?	-85.0	?	-83.0	Stopnisce
9.	26.3.2015	?	-84.8	?	-85.0	Stopnisce
10.	26.3.2015	?	-84.0	?	-85.0	Stopnisce
11.	26.3.2015	?	-86.4	?	-85.0	Stopnisce
12.	26.3.2015	?	-85.1	?	-84.2	Stopnisce
13.	26.3.2015	?	-85.4	?	-79.6	Stopnisce
14.	26.3.2015	?	-84.5	?	-81.0	Stopnisce
15.	26.3.2015	?	?	?	-82.4	Stopnisce
16.	26.3.2015	-87.0	-83.4	-79.8	-83.2	HodnikEna

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
17.	26.3.2015	-90.3	-85.0	?	-82.4	HodnikEna
18.	26.3.2015	-86.0	-82.1	?	-82.3	HodnikEna
19.	26.3.2015	-85.0	-84.0	?	-82.2	HodnikEna
20.	26.3.2015	-82.0	?	?	-84.4	HodnikEna
21.	26.3.2015	-82.0	-86.0	?	?	HodnikEna
22.	26.3.2015	?	-82.1	?	-81.4	HodnikEna
23.	26.3.2015	-85.1	-82.0	?	-84.0	HodnikEna
24.	26.3.2015	-85.0	-86.0	?	?	HodnikEna
25.	26.3.2015	-85.2	-85.0	?	-83.7	HodnikEna
26.	26.3.2015	-87.0	-86.7	?	?	HodnikEna
27.	26.3.2015	-85.0	-81.0	?	-85.0	HodnikEna
28.	26.3.2015	-84.7	-85.9	?	-83.9	HodnikEna
29.	26.3.2015	-86.0	-83.9	?	-81.5	HodnikEna
30.	26.3.2015	-88.0	-84.3	?	-83.0	HodnikEna
31.	26.3.2015	-86.9	-87.0	?	-80.6	HodnikEna
32.	26.3.2015	?	-82.2	-79.2	-79.4	HodnikEna
33.	26.3.2015	?	-81.0	?	-83.0	HodnikEna
34.	26.3.2015	?	-86.8	?	-83.3	HodnikEna
35.	26.3.2015	?	-87.0	?	-81.0	HodnikEna
36.	26.3.2015	?	?	?	-80.0	HodnikEna
37.	26.3.2015	?	-86.6	-82.0	-73.7	HodnikEna
38.	26.3.2015	?	-81.0	?	?	HodnikEna
39.	26.3.2015	?	-82.0	?	-75.5	HodnikEna
40.	26.3.2015	?	-85.1	-79.8	-80.2	HodnikEna
41.	26.3.2015	?	-82.6	-78.9	-75.4	HodnikEna
42.	26.3.2015	?	-82.0	-77.7	-75.5	HodnikEna
43.	26.3.2015	?	-78.9	-79.0	-78.9	HodnikEna
44.	26.3.2015	?	-81.4	-79.7	-75.7	HodnikEna
45.	26.3.2015	?	-78.6	-77.8	-81.4	HodnikEna
46.	26.3.2015	?	-81.0	-76.5	-81.0	HodnikEna
47.	26.3.2015	?	-82.7	-78.1	-79.7	HodnikEna
48.	26.3.2015	?	-80.3	-82.0	-81.0	HodnikEna
49.	26.3.2015	?	-82.3	?	-75.5	HodnikDva
50.	26.3.2015	?	-79.7	?	-74.5	HodnikDva
51.	26.3.2015	?	-85.2	?	-69.0	HodnikDva
52.	26.3.2015	?	-80.2	-81.0	-79.2	HodnikDva
53.	26.3.2015	?	-85.0	?	-77.0	HodnikDva

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
54.	26.3.2015	?	-81.6	-79.5	-71.8	HodnikDva
55.	26.3.2015	?	-83.0	?	-77.0	HodnikDva
56.	26.3.2015	?	-79.8	?	-68.9	HodnikDva
57.	26.3.2015	?	-80.5	-79.0	-67.8	HodnikDva
58.	26.3.2015	?	-81.0	?	-67.5	HodnikDva
59.	26.3.2015	?	-82.9	-81.0	-68.8	HodnikDva
60.	26.3.2015	?	-81.8	-81.0	-67.5	HodnikDva
61.	26.3.2015	?	-83.5	?	-65.0	HodnikDva
62.	26.3.2015	?	-84.5	-81.0	-72.4	HodnikDva
63.	26.3.2015	?	-83.0	?	-72.0	HodnikDva
64.	26.3.2015	?	-83.1	?	-71.1	HodnikDva
65.	26.3.2015	?	-82.3	?	-74.2	HodnikDva
66.	26.3.2015	?	-77.1	?	-70.9	HodnikDva
67.	26.3.2015	?	-80.1	-79.3	-76.7	HodnikDva
68.	26.3.2015	?	-82.0	-79.0	-76.0	HodnikDva
69.	26.3.2015	?	-76.4	-79.5	-80.1	HodnikDva
70.	26.3.2015	?	-77.5	-80.0	-78.0	HodnikDva
71.	26.3.2015	?	-81.5	-77.7	-75.4	HodnikDva
72.	26.3.2015	?	-83.0	-77.7	-78.3	HodnikDva
73.	26.3.2015	?	-78.8	-78.9	-69.0	HodnikDva
74.	26.3.2015	?	-84.3	?	-77.7	HodnikDva
75.	26.3.2015	?	?	-79.0	-71.0	HodnikTri
76.	26.3.2015	?	-82.0	?	-68.4	HodnikTri
77.	26.3.2015	?	-84.0	?	-66.5	HodnikTri
78.	26.3.2015	?	-81.6	?	-55.6	HodnikTri
79.	26.3.2015	?	-81.7	-81.0	-56.2	HodnikTri
80.	26.3.2015	?	-80.4	?	-58.9	HodnikTri
81.	26.3.2015	?	-83.0	?	-60.3	HodnikTri
82.	26.3.2015	?	-86.2	?	-68.9	HodnikTri
83.	26.3.2015	?	-84.0	?	-69.5	HodnikTri
84.	26.3.2015	?	-84.7	-79.1	-69.5	HodnikTri
85.	26.3.2015	?	-83.0	?	-68.0	HodnikTri
86.	26.3.2015	?	-81.4	?	-68.9	HodnikTri
87.	26.3.2015	?	-87.0	?	-67.5	HodnikTri
88.	26.3.2015	?	-83.2	?	-65.8	HodnikTri
89.	26.3.2015	?	-82.3	?	-66.5	HodnikTri
90.	26.3.2015	?	-84.8	?	-64.0	HodnikTri

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
91.	26.3.2015	?	-84.0	?	-63.0	HodnikTri
92.	26.3.2015	?	-82.0	-80.0	-64.4	HodnikTri
93.	26.3.2015	?	-85.0	?	-65.0	HodnikTri
94.	26.3.2015	?	-80.3	?	-62.0	HodnikTri
95.	26.3.2015	?	-84.3	?	-61.5	HodnikTri
96.	26.3.2015	?	-82.7	?	-65.1	HodnikTri
97.	26.3.2015	?	-79.0	?	-65.0	HodnikTri
98.	26.3.2015	?	-84.0	-80.0	-63.0	HodnikTri
99.	26.3.2015	?	-82.7	?	-63.5	HodnikTri
100.	26.3.2015	?	-82.5	?	-58.0	HodnikTri
101.	26.3.2015	?	-81.6	-79.7	-61.1	HodnikTri
102.	26.3.2015	?	-84.3	?	-61.7	HodnikTri
103.	26.3.2015	?	-69.5	?	?	HodnikStiri
104.	26.3.2015	?	-74.8	-75.9	-82.3	HodnikStiri
105.	26.3.2015	?	-59.4	-78.1	-83.3	HodnikStiri
106.	26.3.2015	?	-64.0	-77.0	-83.0	HodnikStiri
107.	26.3.2015	?	-58.4	-79.5	-86.3	HodnikStiri
108.	26.3.2015	?	-57.8	-79.8	-81.0	HodnikStiri
109.	26.3.2015	?	-69.4	-80.3	-82.8	HodnikStiri
110.	26.3.2015	?	-66.3	-78.4	-81.0	HodnikStiri
111.	26.3.2015	?	-69.0	-77.0	-84.0	HodnikStiri
112.	26.3.2015	?	-66.0	-80.0	-82.4	HodnikStiri
113.	26.3.2015	?	-66.0	?	?	HodnikStiri
114.	26.3.2015	?	-64.1	-79.4	-82.8	HodnikStiri
115.	26.3.2015	?	-68.0	?	-84.0	HodnikStiri
116.	26.3.2015	?	-65.1	-78.9	-84.0	HodnikStiri
117.	26.3.2015	?	-66.0	-81.0	-85.0	HodnikStiri
118.	26.3.2015	?	-76.6	-79.4	?	HodnikStiri
119.	26.3.2015	-84.6	?	?	?	P1
120.	26.3.2015	-82.5	?	?	?	P1
121.	26.3.2015	-78.9	?	?	?	P1
122.	26.3.2015	-84.4	?	?	?	P1
123.	26.3.2015	-86.5	?	?	?	P1
124.	26.3.2015	-84.3	?	?	?	P1
125.	26.3.2015	-86.0	?	?	?	P1
126.	26.3.2015	-84.7	?	?	?	P1
127.	26.3.2015	-82.9	?	?	?	P1

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
128.	26.3.2015	-83.0	?	?	?	P1
129.	26.3.2015	-89.0	?	?	?	P1
130.	26.3.2015	-80.9	?	?	?	P1
131.	26.3.2015	-78.5	?	?	?	P1
132.	26.3.2015	-84.3	?	?	?	P1
133.	26.3.2015	-73.6	?	?	?	P1
134.	26.3.2015	-70.8	?	?	?	P1
135.	26.3.2015	-75.1	?	?	?	P1
136.	26.3.2015	-77.4	?	?	?	P1
137.	26.3.2015	-85.0	?	?	?	P1
138.	13.4.2015	-77.0	?	?	?	P1
139.	13.4.2015	-81.0	?	?	?	P1
140.	13.4.2015	-85.0	?	?	?	P1
141.	13.4.2015	-83.0	?	?	?	P1
142.	13.4.2015	-89.0	?	?	?	P1
143.	13.4.2015	-87.0	?	?	?	P1
144.	13.4.2015	-83.0	?	?	?	P1
145.	13.4.2015	-84.0	?	?	?	P1
146.	13.4.2015	-86.0	?	?	?	P1
147.	13.4.2015	-82.0	?	?	?	P1
148.	13.4.2015	-82.0	?	?	?	P1
149.	13.4.2015	-74.0	?	?	?	P1
150.	13.4.2015	-71.0	?	?	?	P1
151.	13.4.2015	-73.0	?	?	?	P1
152.	13.4.2015	-85.0	?	?	?	P1
153.	13.4.2015	-80.0	?	?	?	P1
154.	13.4.2015	-85.0	?	?	?	P1
155.	13.4.2015	-78.0	?	?	?	P1
156.	13.4.2015	-83.0	?	?	?	P1
157.	13.4.2015	-83.0	?	?	?	P1
158.	13.4.2015	-85.0	?	?	?	P1
159.	13.4.2015	-79.0	?	?	?	P1
160.	13.4.2015	-75.0	?	?	?	P1
161.	13.4.2015	-76.0	?	?	?	P1
162.	13.4.2015	-75.0	?	?	?	P1
163.	13.4.2015	-76.8	?	?	?	P1
164.	13.4.2015	-82.0	?	?	?	P1

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
165.	13.4.2015	-75.9	?	?	?	P1
166.	13.4.2015	-77.4	?	?	?	P1
167.	13.4.2015	-83.5	?	?	?	P1
168.	13.4.2015	-81.8	?	?	?	P1
169.	13.4.2015	-81.0	?	?	?	P1
170.	13.4.2015	-79.6	?	?	?	P1
171.	13.4.2015	-81.0	?	?	?	P1
172.	13.4.2015	-82.7	?	?	?	P1
173.	13.4.2015	-82.0	?	?	?	P1
174.	13.4.2015	-78.4	?	?	?	P1
175.	13.4.2015	-83.0	?	?	?	P1
176.	13.4.2015	-78.6	?	?	?	P1
177.	13.4.2015	-82.1	?	?	?	P1
178.	13.4.2015	-83.0	?	?	?	P1
179.	13.4.2015	-79.8	?	?	?	P1
180.	13.4.2015	-77.0	?	?	?	P1
181.	13.4.2015	-79.0	?	?	?	P1
182.	13.4.2015	-83.0	?	?	?	P1
183.	13.4.2015	-83.0	?	?	?	P1
184.	13.4.2015	-81.8	?	?	?	P1
185.	13.4.2015	-83.0	?	?	?	P1
186.	13.4.2015	-86.0	?	?	?	P1
187.	13.4.2015	-80.0	?	?	?	P1
188.	13.4.2015	-78.0	?	?	?	P1
189.	13.4.2015	-79.0	?	?	?	P1
190.	13.4.2015	-77.0	?	?	?	P1
191.	13.4.2015	-80.0	?	?	?	P1
192.	13.4.2015	-87.0	-78.5	?	-80.0	HodnikEna
193.	13.4.2015	-85.0	-80.2	?	-77.7	HodnikEna
194.	13.4.2015	-88.0	-81.3	-79.0	-76.7	HodnikEna
195.	13.4.2015	?	-82.0	?	-77.0	HodnikEna
196.	13.4.2015	?	-77.4	?	-77.8	HodnikEna
197.	13.4.2015	-80.0	?	?	?	HodnikEna
198.	13.4.2015	-86.3	-74.8	?	-79.9	HodnikEna
199.	13.4.2015	-81.0	-81.5	?	?	HodnikEna
200.	13.4.2015	-82.5	-85.0	?	-81.0	HodnikEna
201.	13.4.2015	-86.0	-85.0	?	-79.0	HodnikEna

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
202.	13.4.2015	-79.0	-83.9	-79.0	?	HodnikEna
203.	13.4.2015	-83.0	?	?	?	HodnikEna
204.	13.4.2015	-86.0	-82.0	?	?	HodnikEna
205.	13.4.2015	-86.2	-84.7	-79.0	-79.2	HodnikEna
206.	13.4.2015	-83.0	-81.0	?	-79.7	HodnikEna
207.	13.4.2015	-82.2	-82.9	-79.0	-77.1	HodnikEna
208.	13.4.2015	?	-78.9	?	-68.4	HodnikTri
209.	13.4.2015	?	-76.6	?	-67.8	HodnikTri
210.	13.4.2015	?	-82.0	?	-53.0	HodnikTri
211.	13.4.2015	?	-85.0	?	-63.0	HodnikTri
212.	13.4.2015	?	-83.4	?	-67.4	HodnikTri
213.	13.4.2015	?	-83.0	?	-69.0	HodnikTri
214.	13.4.2015	?	-82.0	?	-67.0	HodnikTri
215.	13.4.2015	?	-87.0	?	-62.0	HodnikTri
216.	13.4.2015	?	-82.5	?	-64.0	HodnikTri
217.	13.4.2015	?	-83.0	?	-62.0	HodnikTri
218.	13.4.2015	?	-82.5	-79.3	-67.9	HodnikTri
219.	13.4.2015	?	-79.0	-79.0	-64.0	HodnikTri
220.	13.4.2015	?	-77.0	?	-67.0	HodnikTri
221.	13.4.2015	?	-81.6	-80.3	-66.3	HodnikTri
222.	13.4.2015	?	-79.0	?	-67.0	HodnikTri
223.	13.4.2015	?	-84.0	-79.0	-66.0	HodnikTri
224.	13.4.2015	?	-81.5	-81.0	-71.0	HodnikTri
225.	13.4.2015	?	-82.4	-79.7	-67.4	HodnikTri
226.	13.4.2015	?	-82.6	?	-70.4	HodnikTri
227.	13.4.2015	?	-84.1	?	-67.0	HodnikTri
228.	13.4.2015	?	?	?	-81.0	Stopnisce
229.	13.4.2015	?	?	?	-79.0	Stopnisce
230.	13.4.2015	?	?	?	-80.0	Stopnisce
231.	13.4.2015	?	?	?	-83.0	Stopnisce
232.	13.4.2015	?	?	?	-79.7	Stopnisce
233.	13.4.2015	-88.0	-82.6	-78.6	-74.5	Stopnisce
234.	13.4.2015	?	-84.4	?	-78.2	Stopnisce
235.	13.4.2015	?	-85.7	?	-78.0	Stopnisce
236.	13.4.2015	?	-85.5	?	-74.9	Stopnisce
237.	13.4.2015	?	-83.7	?	-77.7	Stopnisce
238.	13.4.2015	?	-86.5	?	-79.0	Stopnisce

	Datum	DT_{P1}	DT_{Garaza}	$DT_{Referat}$	DT_{SSFRI}	Regija
239.	13.4.2015	?	-84.6	?	-79.5	Stopnisce
240.	13.4.2015	?	-85.7	?	-79.0	Stopnisce
241.	13.4.2015	?	-79.7	?	-74.3	Stopnisce
242.	13.4.2015	?	-79.5	-79.0	-79.7	Stopnisce